Informatica® Cloud Data Quality
July 2024

Data Quality accelerator bundles

Informatica Cloud Data Quality Data Quality accelerator bundles
July 2024
July 2024

© Copyright Informatica LLC 2022, 2024

This software and documentation are provided only under a separate license agreement containing restrictions on use and disclosure. No part of this document may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording or otherwise) without prior consent of Informatica LLC.

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation is subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License.

Informatica, Informatica Cloud, Informatica Intelligent Cloud Services, and the Informatica logo are trademarks or registered trademarks of Informatica LLC in the United States and many jurisdictions throughout the world. A current list of Informatica trademarks is available on the web at https://www.informatica.com/trademarks.html. Other company and product names may be trade names or trademarks of their respective owners.

Portions of this software and/or documentation are subject to copyright held by third parties.

The information in this documentation is subject to change without notice. If you find any problems in this documentation, report them to us at infa_documentation@informatica.com.

Informatica products are warranted according to the terms and conditions of the agreements under which they are provided. INFORMATICA PROVIDES THE INFORMATION IN THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.

Publication Date: 2024-07-11

Table of Contents

Preface 4

Informatica Resources 4

- Informatica Documentation 4
- Informatica Intelligent Cloud Services web site 4
- Informatica Intelligent Cloud Services Communities 4
- Informatica Intelligent Cloud Services Marketplace 4
- Informatica Knowledge Base 5
- Informatica Intelligent Cloud Services Trust Center 5
- Informatica Global Customer Support 5

Chapter 1: Introduction to Cloud Data Quality accelerator bundles 6

Using bundles 6

- Copying bundle assets 7
- Adding an asset to a mapplet or mapping 7
- Referencing bundle assets 8

Chapter 2: Core Data Quality Bundle 9

Chapter 3: Data Quality bundle for financial services 25

Chapter 4: Data Quality bundle for crisis responses 40

Cloud Data Quality Crisis Response Bundle 40

Chapter 5: Data Quality bundle for Americas 55

Cloud Data Quality Americas bundle 55

Chapter 6: Data Quality bundle for United Kingdom 70

Chapter 7: Data Quality bundle for Australia and New Zealand 84

Chapter 8: Data Quality bundle for Europe 95

Chapter 9: Accelerating data projects for automation 110

Cloud Data Quality automation bundle 110

Index 114

Preface

Read *Data Quality bundles* to learn about the accelerator bundles that Informatica prepares for Informatica Cloud Data Quality users. Accelerator bundles address common data quality issues in a country, a region, or an industry. Accelerators can include prebuilt rules and reference data objects that you can copy or reference in your cloud environment.

Informatica Resources

Informatica provides you with a range of product resources through the Informatica Network and other online portals. Use the resources to get the most from your Informatica products and solutions and to learn from other Informatica users and subject matter experts.

Informatica Documentation

Use the Informatica Documentation Portal to explore an extensive library of documentation for current and recent product releases. To explore the Documentation Portal, visit https://docs.informatica.com.

If you have questions, comments, or ideas about the product documentation, contact the Informatica Documentation team at infa_documentation@informatica.com.

Informatica Intelligent Cloud Services web site

You can access the Informatica Intelligent Cloud Services web site at http://www.informatica.com/cloud.

Informatica Intelligent Cloud Services Communities

Use the Informatica Intelligent Cloud Services Community to discuss and resolve technical issues. You can also find technical tips, documentation updates, and answers to frequently asked questions.

Access the Informatica Intelligent Cloud Services Community at:

https://network.informatica.com/community/informatica-network/products/cloud-integration

Developers can learn more and share tips at the Cloud Developer community:

https://network.informatica.com/community/informatica-network/products/cloud-integration/cloud-developers

Informatica Intelligent Cloud Services Marketplace

Visit the Informatica Marketplace to try and buy Data Integration Connectors, templates, and mapplets:

https://marketplace.informatica.com/

Informatica Knowledge Base

Use the Informatica Knowledge Base to find product resources such as how-to articles, best practices, video tutorials, and answers to frequently asked questions.

To search the Knowledge Base, visit https://search.informatica.com. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team at KB_Feedback@informatica.com.

Informatica Intelligent Cloud Services Trust Center

The Informatica Intelligent Cloud Services Trust Center provides information about Informatica security policies and real-time system availability.

You can access the trust center at https://www.informatica.com/trust-center.html.

Subscribe to the Informatica Intelligent Cloud Services Trust Center to receive upgrade, maintenance, and incident notifications. The Informatica Intelligent Cloud Services Status page displays the production status of all the Informatica cloud products. All maintenance updates are posted to this page, and during an outage, it will have the most current information. To ensure you are notified of updates and outages, you can subscribe to receive updates for a single component or all Informatica Intelligent Cloud Services components. Subscribing to all components is the best way to be certain you never miss an update.

To subscribe, on the Informatica Intelligent Cloud Services Status page, click **SUBSCRIBE TO UPDATES**. You can choose to receive notifications sent as emails, SMS text messages, webhooks, RSS feeds, or any combination of the four.

Informatica Global Customer Support

You can contact a Global Support Center through the Informatica Network or by telephone.

To find online support resources on the Informatica Network, click **Contact Support** in the Informatica Intelligent Cloud Services Help menu to go to the **Cloud Support** page. The **Cloud Support** page includes system status information and community discussions. Log in to Informatica Network and click **Need Help** to find additional resources and to contact Informatica Global Customer Support through email.

The telephone numbers for Informatica Global Customer Support are available from the Informatica web site at https://www.informatica.com/services-and-training/support-services/contact-us.html.

CHAPTER 1

Introduction to Cloud Data Quality accelerator bundles

A bundle is a set of Informatica Intelligent Cloud Services assets that you can use to improve productivity and enhance the quality of your data. A data quality accelerator is bundle of assets that address common data quality problems in a country, a region, or an industry.

The **Data Quality Core Bundle** is a set of pre-configured data quality assets that can parse, standardize, validate, and, deduplicate business data. The assets provide solutions for general data, address data, contact data, corporate data, and product data.

The accelerator may contain a combination of individual assets, mapplet rules, and dictionaries. Add the assets to your mappings to accelerate the deployment of data quality solutions in your organization. Before you use an accelerator bundle, copy it to a project so that you can customize the assets for your business objectives.

You can find accelerator bundles for your organization on the **Explore** page.
Informatica provides a core accelerator that can help you to quickly develop a range of data quality solutions for general business data issues. Informatica also provides accelerators specifically for American organizations, for the rapid deployment of data projects in crisis response scenarios, and to support the automation of data profiling.

Using bundles

Use the assets in a bundle when the assets are designed to meet the same objectives as your project.

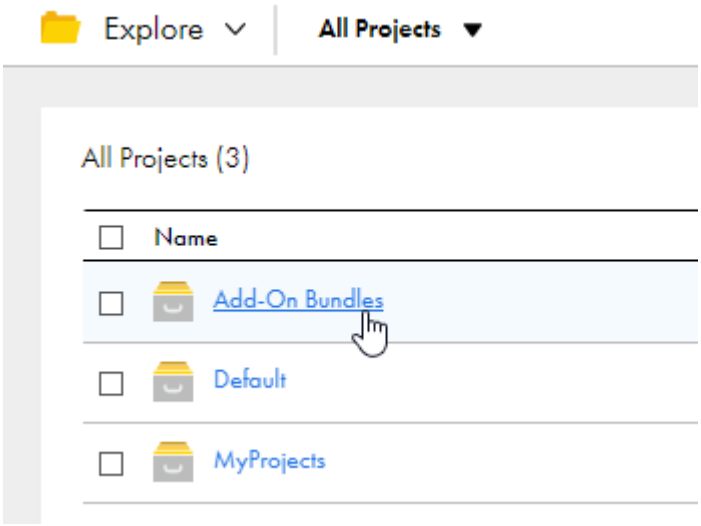
How you can include bundle assets in your project depends on the permissions that the bundle designer designates for the bundle. You can perform one of the following actions depending on which option the bundle designer allows:

- Copy the bundle assets to your project folder and edit the assets for your project. Bundles that allow copy only can be located in any project folder.
- Reference the bundle assets. Select the assets from the Add-on Bundles project when you configure a task. You cannot change a bundle that only allows reference. Bundles that allow reference only are always located in the Add-on Bundles project.
- Copy or reference the bundle assets. Either copy the bundle assets to your project folder or reference the bundle assets in the Add-on Bundles project. Bundles that allow reference or copy are always located in the Add-on Bundles project.

Copying bundle assets

If a bundle allows the copy action, you can copy all of the assets in a bundle to a project or folder.

1. Open the **Data Quality** service or the **Data Integration** service.



2. On the **Explore** page, create or verify a project that you'll use as a location for the bundle that you select.

   You can copy a bundle directly to the project or to a folder in the project.
3. Open the **Administrator** service.
4. Select **Add-On Bundles** from the left-hand pane, and select the **Available Bundles** tab.

   The tab lists the bundles that you can copy to your project.
5. Select an accelerator bundle to copy.

   The bundle home page opens.
6. Click **Copy Bundle Content To...**
7. Browse to the project or project folder that you identified earlier, and click **Select**.

   If you copy a bundle directly to the project level, the bundle contents appear at the project level and not in a folder within the project.

Note: Informatica updates the contents of accelerator bundles in Informatica Intelligent Cloud Services from time to time. The updates appear in the bundles that you find under **Add-On Bundles** in the Administrator service. The updates do not affect any copy of a bundle that you previously made.

Adding an asset to a mapplet or mapping

After you copy an accelerator to a project on the **Explore** page, you can add an asset in the accelerator to a mapplet or mapping in **Data Integration**.

You can add a cleanse, deduplicate, labeler, parse, rule specification, or verifier asset to a transformation in a mapping. Or, you can add a mapplet asset to a mapping.

1. Open or create a mapping in **Data Integration**.

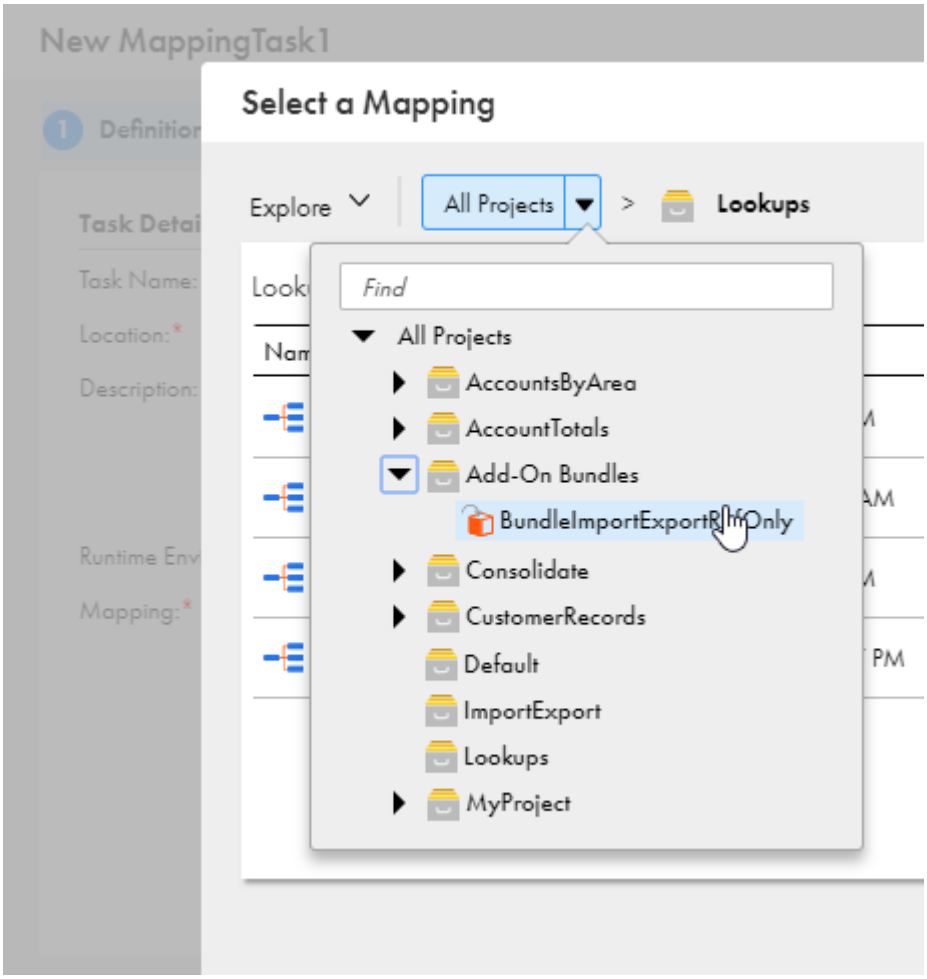
2. Click **Select** in the asset that you're working on, and find the project that contains the asset that you need.
   a. To add a single asset to a transformation, open the transformation and use the asset selection options to find the asset.
   b. To add a mapplet to a mapping, open the mapplet in the mapping and use the mapplet selection options to find the mapplet.
3. Click the row that contains the asset, and then click **Select**.

Referencing bundle assets

Certain asset types allow you to select an asset to use in the asset. If a bundle allows the reference action, you can select the assets in the bundle to use in your asset.

Add a reference to a bundle asset when you configure a mapplet or a mapping in Data Integration.

1. Click **Select** in the asset that you're working on, and then navigate to the Add-On Bundles folder.



2. Open the bundle folder that contains the asset you want to use.
3. Click the row that contains the asset, and then click **Select**.

CHAPTER 2

Core Data Quality Bundle

Informatica provides a core accelerator bundle that can help you to quickly develop a range of data quality solutions for general business data issues.

The following table lists the assets contained in the Data Quality core bundle:

Name	Asset Type	Description
bs_email_EU	BUSINESS SERVICE	Uses Europe based endpoint to access email verification algorithm that identifies valid and non-valid email addresses.
bs_email_US	BUSINESS SERVICE	Uses United States based endpoint to access email verification algorithm that identifies valid and non-valid email addresses.
bs_telephone_EU	BUSINESS SERVICE	Global Phone Validation is a data validation and enrichment service for telephone numbers throughout the world, accessed through an Europe based endpoint.
bs_telephone_US	BUSINESS SERVICE	Global Phone Validation is a data validation and enrichment service for telephone numbers throughout the world, accessed through a United States based endpoint.
c_Add_Space_Around_Ampersand	CLEANSE	Inserts a character space on either side of an ampersand, and reduces multiple spaces to a single space.
c_Add_Space_Around_Hyphen	CLEANSE	Inserts a character space on either side of a hyphen, and reduces multiple spaces to a single space.
c_Add_Space_Around_Period	CLEANSE	Inserts a character space on either side of a period, and removes multiple spaces.
c_Assign_DQ_Match_Code_Description	CLEANSE	Returns descriptions for a Match Code value that a verifier asset generates.
c_av6_address_type_desc	CLEANSE	Returns descriptions for the Address Type Codes.
c_av6_result_quality_desc	CLEANSE	Returns descriptions for Result Quality.

Name	Asset Type	Description
c_av6_verification_status_code_desc	CLEANSE	Returns descriptions for verification status codes.
c_concatenate_words	CLEANSE	Concatenates two words with a space.
c_Country_ISO2Char	CLEANSE	Standardizes country names to their respective two-letter ISO country codes.
c_Country_ISO3Char	CLEANSE	Standardizes country names to their respective three-letter ISO country codes.
c_dq_Label_AlphaNumeric	CLEANSE	Assigns labels to alphanumeric values.
c_ISO_Full_Country_Name	CLEANSE	Returns ISO country name for any country identifier in an input field.
c_Label_Default_Date_Patterns	CLEANSE	Identifies default or sample date values such as 00/00/0000.
c_Label_Default_Numeric_Values	CLEANSE	Identifies default or sample numeric values such as 0 or 999.
c_Label_Default_Text_Values	CLEANSE	Identifies default or sample alphabetic values such as XXX, NA and Default.
c_Lowercase	CLEANSE	Converts the input text to lowercase.
c_Merge_Parse_Number	CLEANSE	Merges the Remainder, Unparsed and Overflowed columns from the parser.
c_Merge_Parse_Text	CLEANSE	Merges the Before, After, Overflow and Unparsed columns from the parser. Removes leading and trailing spaces.
c_merge_remaining_text	CLEANSE	Merges multiple input fields into a single field.
c_Merge_VerificationStatus	CLEANSE	Merges two ports together to create one output port
c_Remove_Apostrophe	CLEANSE	Removes all apostrophe characters from the input.
c_Remove_Extra_Spaces	CLEANSE	Replaces multiple consecutive spaces with a single space and trims leading and trailing spaces.
c_Remove_Hyphen	CLEANSE	Removes all hyphens from an input string.
c_Remove_Leading_Zero	CLEANSE	Removes zero at the start of an input string.
c_Remove_Limited_Punctuation	CLEANSE	Removes any forward slash, backslash, exclamation point, period, or underscore character from an input string.

Name	Asset Type	Description
c_Remove_Parentheses	CLEANSE	Removes all occurrences of left and right parentheses from a string.
c_Remove_Period	CLEANSE	Removes a period from a string.
c_Remove_Period_Parentheses	CLEANSE	Removes all occurrences of a period and left and right parentheses in a string.
c_Remove_Punctuation	CLEANSE	Removes all occurrences of punctuation from the input field.
c_Remove_Quotation	CLEANSE	Removes all double quotes from a string of text.
c_Remove_Slashes	CLEANSE	Removes all occurrences of forward and back slashes in a string.
c_Remove_Space	CLEANSE	Removes all occurrences of a character space from the input field.
c_remove_spaces	CLEANSE	Removes all spaces from an input string.
c_Remove_Space_Join	CLEANSE	Merges two ports and removes any spaces around and between.
c_Replace_Amp_with_Space	CLEANSE	Replaces an ampersand character with a character space.
c_Replace_Back_Slash_with_Fwd_Slash	CLEANSE	Replaces forward-slash with backward-slash.
c_Replace_Diacritic_Chrs	CLEANSE	Replaces letters that use diacritic symbols with their English-language counterparts.
c_Replace_Hyp_Underscore_with_Space	CLEANSE	Replaces a hyphen or underscore character with a space.
c_Replace_Hyph_w_Space	CLEANSE	Replaces a hyphen with a space and reduces multiple spaces to a single space.
c_Replace_Limited_Punctuation_w_Space	CLEANSE	Replaces any forward slash, backslash, exclamation point, period, or underscore character with a character space. Also replaces instances of multiple character spaces with a single space.
c_Replace_NonAlpha_with_Space	CLEANSE	Replaces a number or other non-alphabetic character with a space, and reduces multiple spaces to a single space.
c_Replace_Period_with_Space	CLEANSE	Replaces a period with a character space.
c_Replace_Slashes_w_Space	CLEANSE	Replaces any slash symbol with a character space.

Name	Asset Type	Description
c_Titlecase	CLEANSE	Converts the input text to title case with dictionary-defined exceptions.
c_TrimData	CLEANSE	Trims leading and trailing spaces from the input string.
c_Uppercase	CLEANSE	Converts the input text to uppercase.
c_Validation_Domain	CLEANSE	Validates the domain of an email.
character_sets	DICTIONARY	Contains a set of characters (A-Z , a-z , 0-9).
diacritics_translate_infa	DICTIONARY	Contains a list of characters that use diacritics symbols.
domain_names_infa	DICTIONARY	Contains a list of domain names.
dq_av6_address_type_desc	DICTIONARY	Contains a list of result quality codes with descriptions.
dq_av_match_code_descriptions_infa	DICTIONARY	Contains alphanumeric status values that indicate the outcome of the verification operation for an address. Each status value has a corresponding text description.
dq_av6_result_quality_descriptions	DICTIONARY	Contains a list of the result quality codes that Address Verification version 6 engine returns with descriptions.
dq_av6_status_code_descriptions_infa	DICTIONARY	Contains alpha status values that indicate the outcome of the verification operation for an address. Each status value has a corresponding text description.
iso_full_country_name_infa	DICTIONARY	Contains the short and long forms of country names and their respective two-letter and three-letter ISO country codes.
numeric_default_values_infa	DICTIONARY	Assigns numeric values.
string_default_values_infa	DICTIONARY	Contains a list of sample or default values.
suspect_names_infa	DICTIONARY	Contains a list of suspect names.
user_defined_dates_infa	DICTIONARY	Contains a set of valid calendar date formats.
cities_of_world_infa	DICTIONARY	Contains the cities of the world.
color_base_infa	DICTIONARY	Contains the colors of various shades.
colors_infa	DICTIONARY	Contains a list of color names and their alternatives.

Name	Asset Type	Description
currency_codes_country_codes_infa	DICTIONARY	Contains a list of currency and country codes.
currency_codes_infa	DICTIONARY	Contains the standardized currency codes of all the currencies in the world.
currency_symbol_from_code_infa	DICTIONARY	Contains a list of currency symbols for the respective currency.
date_month_infa	DICTIONARY	Contains the number of the month.
day_month_infa	DICTIONARY	Contains the standard DD/MM format for the days and months of the year.
english_alphabets_infa	DICTIONARY	Contains a list of English letters.
english_consonants_infa	DICTIONARY	Contains the consonants of the English language.
household_infa	DICTIONARY	Contains a list of household elements.
iso_2chr_country_code_infa	DICTIONARY	Contains the two-character ISO country code and the corresponding country name.
iso_3chr_country_code_infa	DICTIONARY	Contains the three-character ISO country code and the corresponding country name.
job_position_infa	DICTIONARY	Contains the general list of job positions.
lbl_Chars	LABELER	Assigns labels to alphabetical characters, spaces and non-alphabetical characters.
lbl_Non_Alpha_Characters	LABELER	Replaces a number or other non-alphabetical characters with 9/S.
major_cities_cntries_infa	DICTIONARY	Contains the list of major cities.
mplt_Comparison_EditDistance	MAPPLET	Calculates the minimum cost of transforming one string to another through insertion, deletion, or replacement of characters.
mplt_Comparison_HammingDistance	MAPPLET	Calculates a match score for two strings by computing the number of positions in which characters differ between the strings. For strings of different length, each additional character in the longest string is counted as a difference between the strings.
mplt_Comparison_JaroDistance	MAPPLET	Reflects the degree of similarity between the first four characters of both the strings and the number of identified character transpositions.

Name	Asset Type	Description
mplt_Comparison_ReverseHammingDistance	MAPPLET	Calculates a match score for two strings by computing the number of positions in which characters differ between the strings. For strings of different length, the algorithm counts each additional character in the longest string as the difference between the strings.
month_to_mm_infa	DICTIONARY	Contains the integer value of the corresponding month of the year.
months_to_num_infa	DICTIONARY	Contains the number of the month in a year. For example, 01 for January, 02 for February.
NAICS_2012_codes_desc_infa	DICTIONARY	Contains version 2012 of the North American Industry Classification System (NAICS) codes and descriptions.
NAICS_2017_codes_desc_infa	DICTIONARY	Contains version 2017 of the North American Industry Classification System (NAICS) codes and descriptions.
NAICS_2022_codes_desc_infa	DICTIONARY	Contains version 2022 of the North American Industry Classification System (NAICS) codes and descriptions.
NAICS_2022_description_infa	DICTIONARY	Contains version 2022 of the North American Industry Classification System (NAICS) codes and descriptions with NAICS description as a valid column.
NAICS_to_SIC_infa	DICTIONARY	Contains NAICS to SIC codes crosswalk.
ProcessStatusMap_V6_V5	DICTIONARY	Contains lists of process status values from version 5 and version 6 of the address verification engine.
profanity_infa	DICTIONARY	Contains a list of profane words.
quantity_list_infa	DICTIONARY	Contains a list of quantities. For example, length, mass, time, area, acceleration.
races_infa	DICTIONARY	Contains the different races of the world.
religion_infa	DICTIONARY	Contains a list of religions.
rs_IsNumeric	DICTIONARY	Checks if the input value is numeric or non-numeric.
rs_prepend_Zero_to_Single_Digit	DICTIONARY	Prepends zero to a single digit. If the input data is not a single digit, the data is returned.
rs_remove_all_leading_zeros	DICTIONARY	Removes all the zeros at the beginning of an input string.

Name	Asset Type	Description
SIC_to_NAICS_infa	DICTIONARY	Contains Standard Industrial Classification (SIC) to North American Industry Classification System (NAICS) codes and descriptions. SIC is designated a valid column.
SIC_To_NAICS_2012	DICTIONARY	Contains Standard Industrial Classification (SIC) codes and descriptions and their closest 2012 North American Industry Classification System (NAICS) codes and descriptions. **Note:** There are no direct relationships between the SIC and NAICS codes, some SIC codes have been expanded out to multiple NAICS codes.
SIC_To_NAICS_2017	DICTIONARY	Contains Standard Industrial Classification (SIC) codes and descriptions and their closest 2017 North American Industry Classification System (NAICS) codes and descriptions. **Note:** There are no direct relationships between the SIC and NAICS codes, some SIC codes have been expanded out to multiple NAICS codes.
std_phone_formats_emea_infa	DICTIONARY	Contains the standard phone formats of the EMEA region.
surname_case_excptn_infa	DICTIONARY	Contains the list of surname with common surname prefixes like "Mc" or "Mac".
title_case_excptn_infa	DICTIONARY	Contains list of words that are title case exceptions. The capitalized major words and minor words are in lowercase.
uom_infa	DICTIONARY	Contains the abbreviations of the Unit of Measures (UOM).
international_dialing_codes_infa	DICTIONARY	Returns the International country phone code when the country name or ISO code is passed.
mplt_Add_Leading_Zeroes	MAPPLET	Adds a leading zero to a string input.
mplt_Add_Parentheses_At_Start_End_ofLine	MAPPLET	Adds an open parenthesis at the beginning and a close parenthesis at the end of an input string.
mplt_Add_Plus_To_Start_of_Line	MAPPLET	Inserts a plus sign in front of an input string.
mplt_Add_Space_Around_Ampersand	MAPPLET	Inserts a space on either side of an ampersand, and removes multiple spaces.
mplt_Add_Space_Around_Hyphen	MAPPLET	Ensures that a single space precedes and follows a dash or a hyphen.

Name	Asset Type	Description
mplt_Add_Space_Between_Number_Letter	MAPPLET	Adds a space between a single digit and a single letter. For example, 3G is returned as 3 G. **Note:** The operation applies to the first alphanumeric pattern in the input string, reading from left to right.
mplt_Add_Spaces_Around_Period	MAPPLET	Inserts a character space on either side of a period, and removes multiple spaces.
mplt_AllTrim	MAPPLET	Removes all leading and trailing spaces from an input string.
mplt_Assign_DQ_Match_Code_Description	MAPPLET	Describes the match score value that a verifier asset returns.
mplt_AV6_Assign_DQ_Status_Code_Description	MAPPLET	Describes the Match Score value that a verifier asset returns for the address verification version 6 engine.
mplt_Comparison_Bigram	MAPPLET	Calculates a match score for two data strings based on the occurrence of consecutive characters in both strings. The mapplet looks for pairs of consecutive characters that are common to both strings. It divides the number of pairs that match in both strings by the total number of character pairs.
mplt_Comparison_Dates	MAPPLET	Compares two dates in an input of type string and returns four outputs. The outputs represent the difference between the dates in hours, days, months, and years respectively. The output values are exclusive of each other. Do not use them together to represent the difference between the data values.
mplt_Completeness	MAPPLET	Checks if the input field is empty or null.
mplt_Completeness_Multi_Port	MAPPLET	Checks if the input field is empty or null. The mapplet has ten input and output ports.
mplt_Concatenate_Words	MAPPLET	Concatenates two strings with a character space as the join character.
mplt_convert_V6_ProcessStatus_to_V5	MAPPLET	Converts the process status values that version 5 of the address verification engine generates to their version 6 equivalents.
mplt_Country_Name_Standardization	MAPPLET	Reads a country name as input and returns the full ISO country name, the ISO three-letter country code, and the ISO two-letter country code.
mplt_Date_Complete	MAPPLET	Checks if an input date string matches any user-defined default value. Does not validate the input string as a date.

Name	Asset Type	Description
mplt_Date_Of_Birth_Validation	MAPPLET	Checks the number of years between the input date and the current date. Returns whether the years represent an adult or a minor. Returns Valid for 0 - 120 inclusive and returns Invalid for all other entries.
mplt_Date_Standardization	MAPPLET	Formats an input date string using user-defined variables. Includes the exp_ParseDate and exp_FormatDate transformations. Input dates that are invalid or do not fit the designated input format are written as all zeros.
mplt_Date_Validation	MAPPLET	Validates an input date string. The input date format variable allows the user to specify the input date format.
mplt_Date_Validation_Variable_Format	MAPPLET	Validates an input date string. The input date format variable allows the user to specify the format the input date should be in.
mplt_Days_Between_Dates	MAPPLET	Calculates the difference between the two dates in days. The result is the difference between the end date and the start date. Input data type should be set as String.
mplt_Days_From_Current_Date	MAPPLET	Calculates the number of days between an input date and the current date (system date). Sets the expected input type as String.
mplt_Email_Parse	MAPPLET	Parses an email from within a string of text.
mplt_Email_Validation	MAPPLET	Validates the email format. It does not check if the email is an activated address. Returns Valid or Invalid.
mplt_email_validation_webservice_EU	MAPPLET	Calls the email web service to process emails. **Note:** Users must update the expression asset with their user ID and password or license key.

Name	Asset Type	Description
mplt_Geocoordinate_In_Polygon	MAPPLET	Determines if a coordinate point falls inside or outside a set of coordinates that form a polygon shape. Enter the latitude and longitude values for the coordinate point in the Coordinate_Point field. Enter the latitude and longitude values that define the polygon in the Polygon_Coordinates field. Enter the latitude and longitude values in decimal format. Use a semi-colon to separate the values. For example, `17.25458; 78.244393` When you enter the polygon coordinates, use a pipe symbol to separate each set of latitude and longitude values. For example, `88;-179.999999| 89.999999;-179.999999| 89.999999;179.999999|88;179.999999` The polygon coordinates must connect to each other in a sequence that describes the perimeter of the polygon. For example, the first point connects to the second point, and the second point connects to the third point.
mplt_Global_Address_Validation_Hybrid	MAPPLET	Validates the deliverability of address records from multiple countries. The mapplet corrects errors where possible. Use the rule with input address fields in a hybrid format.
mplt_Identify_Suspect_Names	MAPPLET	Flags suspicious or fake names in input data.
mplt_IsNumeric	MAPPLET	Checks if input data is numeric.
mplt_LowerCase	MAPPLET	Converts input text to lowercase characters.
mplt_Negative_Number_Validation	MAPPLET	Checks if an input is a number and less than zero.
mplt_Numeric_Completeness	MAPPLET	Checks the completeness of a number field. Checks that the field contains a number, is not null or blank, and does not match a sample or default value.
mplt_Parse_Alpha_Chars_from_Non_Alpha_Chars	MAPPLET	Parses alphabetic characters and spaces from non-alphabetic characters.
mplt_Parse_First_Word	MAPPLET	Parses the first text string from an input field, reading from left to right. Uses a character space to delimit text strings.
mplt_Parse_Number_At_End_Of_Line	MAPPLET	Parses a number that occurs at the end of an input string, reading from left to right.
mplt_Parse_Number_At_Start_Of_Line	MAPPLET	Parses a number that occurs at the beginning of a string.

Name	Asset Type	Description
mplt_Parse_Text_Between_Parentheses	MAPPLET	Parses a text string between left and right parentheses from an input field.
mplt_Parse_Text_in_Single_Quotes	MAPPLET	Parses a text string between single quotes from an input field. If there are multiple instances, the final occurrence in the field is parsed out.
mplt_Past_Date_Label	MAPPLET	Flags a past date as invalid, and converts a future date to the current system date.
mplt_Prepend_Zero_to_Single_digit	MAPPLET	Prepends a zero to a single digit. If the input is not a single digit, returns the input data.
mplt_Remove_All_Leading_Zeros	MAPPLET	Removes any leading zero from number data.
mplt_Remove_Apostrophe	MAPPLET	Removes apostrophes from a string.
mplt_Remove_Control_Characters	MAPPLET	Identifies and removes non-printable characters in an input field.
mplt_Remove_Extra_Spaces	MAPPLET	Replaces multiple consecutive character spaces with a single space, and trims leading and trailing spaces.
mplt_Remove_Hyphen	MAPPLET	Removes all hyphens from an input string.
mplt_Remove_Leading_Zero	MAPPLET	Removes a leading zero from a string.
mplt_Remove_Limited_Punctuation	MAPPLET	Removes a limited number of punctuation symbols, including a forward slash, backslash, exclamation point, period, or underscore character. Also replaces instances of two, three, or four consecutive character spaces with a single space.
mplt_Remove_Non_Numbers	MAPPLET	Removes all characters that are not a digit or number from an input string.
mplt_Remove_Parentheses	MAPPLET	Removes all occurrences of left and right parentheses from a string.
mplt_Remove_Period	MAPPLET	Removes all occurrences of a period from a string.
mplt_Remove_Period_Parantheses	MAPPLET	Removes all occurrences of a period and left and right parentheses in a string.
mplt_Remove_Punctuation	MAPPLET	Removes all punctuation symbols from the input.
mplt_Remove_Punctuation_and_Space	MAPPLET	Removes all punctuation symbols and character spaces from an input field, and returns the remaining digits and letters.

Name	Asset Type	Description
mplt_Remove_Quotation	MAPPLET	Removes all double quotation marks from a string of texts. Does not replace with character spaces.
mplt_Remove_Slashes	MAPPLET	Removes all occurrences of forward slashes and backslashes in a string of text.
mplt_Remove_Space	MAPPLET	Removes all occurrences of a character space from an input field.
mplt_Replace_Ampersand_with_Space	MAPPLET	Replaces an ampersand symbol with a character space.
mplt_Replace_Hyphen_Underscore_with_Space	MAPPLET	Replaces forward slash or backslash symbols with a character space.
mplt_Replace_Hyphen_with_Space	MAPPLET	Replaces any occurrence of a hyphen with a single space.
mplt_Replace_Limited_Punct_with_Space	MAPPLET	Replaces punctuation symbols including a forward slash, backslash, exclamation point, period, or underscore character with a character space. Also replaces instances of two, three, or four consecutive character spaces with a single space.
mplt_Replace_Non_Alphabetic_with_Space	MAPPLET	Replaces all non-alphabetic characters with a character space. Trims multiple successive spaces to a single space.
mplt_Replace_Period_With_Space	MAPPLET	Replaces any occurrence of a period with a character space.
mplt_Replace_Punctuation_with_Space	MAPPLET	Replaces punctuation symbols with a character space.
mplt_Replace_Slashes_with_Space	MAPPLET	Replaces all occurrences of a forward slash or backslash with a character space.
mplt_Reverse_String_Input	MAPPLET	Reverses the order of the characters in an input field.
mplt_String_Completeness	MAPPLET	Checks if an input field is empty, null, or contains sample or default values.
mplt_TitleCase	MAPPLET	Converts input text to title case, with dictionary-defined exceptions.
mplt_UpperCase	MAPPLET	Changes alphabetic characters to uppercase.
mplt_Verhoeff_Algorithm	MAPPLET	Validates the check digit of a number. Returns valid if the check digit (the final digit of the input) is correct. Returns invalid for incorrect check digits and for all non-numeric input.

Name	Asset Type	Description
mplt_Years_Since_Date_Of_Birth	MAPPLET	Returns the number of years since a given date.
mplt_Email_Parse_and_Validate	MAPPLET	Parses an email from a string of texts, provides the elements of mailbox, domain and sub-domain, and validates if the structure and domain are valid. It does not check if the email is an activated address.
mplt_Email_Parse_Into_Mailbox_Domain	MAPPLET	Parses an email into its constituent elements of mailbox, domain and sub-domain. For example, `info@informatica.com` will be parsed into `info` (mailbox), `informatica` (sub-domain), and `com` (domain). The mapplet assumes that the text being entered is a valid email.
mplt_email_validation_webservice_US	MAPPLET	Mapplet calls the Email Web Service to process emails. Email Verification combines state-of-the-art email verification algorithms with hygiene algorithms to allow customers to detect and avoid dangerous or potentially dangerous email addresses or domains. You must update the expression asset to include their user IDs and passwords or license keys.
mplt_telephone_webservice_EU	MAPPLET	Mapplet calls the Telephone Web Service to process global telephone numbers. Global Phone Validation is a data validation and enrichment service for telephone numbers throughout the world. You must update the expression asset to include their user IDs and passwords or license keys.
mplt_telephone_webservice_US	MAPPLET	Mapplet calls the Telephone Web Service to process global telephone numbers. Global Phone Validation is a data validation and enrichment service for telephone numbers throughout the world. You must update the expression asset to include their user IDs and passwords or license keys.
mplt_Translate_Diacritic_Characters	MAPPLET	Translates diacitical characters to regular ascii text characters. For example, it changes an `ë` to an `e` (returns `tëst` as `test`).
p_control_characters	PARSE	Identify and parse control characters from input string.
p_Email_from_Text	PARSE	Parses an email address from a text string.
p_First_Word	PARSE	Parses the first word of a string.

Name	Asset Type	Description
p_Number_At_Start_of_String	PARSE	Parses a number at the beginning of a string.
p_Parse_Number_End_of_Line	PARSE	Parses a number that occurs at the end of a string.
p_Suspect_Names	PARSE	Parses suspicious names.
p_Text_btw_Parentheses	PARSE	Parses text between parentheses.
p_mailbox_subdomain	PARSE	Parses the first word (sub-domain) of an email address.
p_top_Level_domain	PARSE	Parses top level domain of an email address.
p_Text_btw_Single_Quotes	PARSE	Parses text between single quotes.
Data Quality Core Bundle	PROJECT	Covers general, address, contact, corporate and product data. Copy the content of the bundle to a project or folder. Use the copies or update them as per the business requirement of the user. This bundle is updated from time to time, and modifications will not affect the changes that users make in the copied assets.
rs_check_input_is_not_null	RULE SPECIFICATION	Verifies if the input is null.
rs_convert_integers_to_date_time	RULE SPECIFICATION	Takes the date input fields and combines them to create a date that consists of year, month, day, hour, minute, second, nanosecond. Minimum input required is year, month, day.
rs_convert_string_input_to_date_format	RULE SPECIFICATION	Takes an input string in the mm/dd/yyyy format and converts it to the date/time format.
rs_convert_to_lower_case	RULE SPECIFICATION	Converts input strings to lowercase.
rs_convert_to_sentence_case	RULE SPECIFICATION	Converts input strings to sentence case.
rs_convert_to_title_case	RULE SPECIFICATION	Converts input strings to title case.
rs_convert_to_toggle_case	RULE SPECIFICATION	Converts input strings to toggle case.
rs_convert_to_upper_case	RULE SPECIFICATION	Converts input strings to uppercase.
rs_extract_year_from_date	RULE SPECIFICATION	Returns the year from a date/time input in the format yyyy-mm-dd or yyyy-MM-dd HH:mm:ss.SSS.

Name	Asset Type	Description
rs_increment_date	RULE SPECIFICATION	Adds a number that you specify to the day, month, or year value in the date.
rs_join_two_inputs_with_dash	RULE SPECIFICATION	Merges two input strings into a single output, with a dash (-) as the join character.
rs_join_two_inputs_with_space	RULE SPECIFICATION	Merges two input strings into a single output, with a character space as the join character.
rs_join_two_strings	RULE SPECIFICATION	Concatenates two strings into a single string with no join character.
rs_null_check	RULE SPECIFICATION	Identifies null values in an input field. Returns `1` for a null value, and returns `0` if the input is not null.
rs_output_contents_of_selected_column	RULE SPECIFICATION	Returns a value from a string that contains multiple delimited values. The value is determined by an integer in the first input that the rule specifies.
rs_remove_custom_strings	RULE SPECIFICATION	Removes an input string that the rule specification logic specifies.
rs_remove_international_dialing_code	RULE SPECIFICATION	Removes international dialing code data from an input string.
rs_remove_leading_trailing_spaces	RULE SPECIFICATION	Removes spaces at the start and end of an input string.
rs_remove_multiple_spaces	RULE SPECIFICATION	Removes multiple character spaces from an input string.
rs_replace_custom_strings	RULE SPECIFICATION	Finds and replaces input strings that the rule specification logic specifies.
rs_return_greatest_across_columns	RULE SPECIFICATION	Sorts a list of values in numerical order or alphabetical order and returns the highest value from the list. The values must be the same data type.
rs_return_lessor_across_columns	RULE SPECIFICATION	Returns the lessor of two values across two input fields on the same data row.
rs_return_string_length	RULE SPECIFICATION	Returns the number of characters in a string.
rs_reverse_string	RULE SPECIFICATION	Returns the characters and data in an input string in reverse order.
rs_Assign_DQ_Match_Code_Description	RULE SPECIFICATION	Returns the text description of a match code value that a verifier asset generates.
rs_ISO_Full_Country_Name	RULE SPECIFICATION	Returns the ISO country name for any country identifier in an input field.

Name	Asset Type	Description
rs_Standardize_Currency_Inputs	RULE SPECIFICATION	Standardizes the currency symbol based on current symbol, three-letter code, or full currency name.
rs_truncate_decimal_value	RULE SPECIFICATION	Truncates a float data value to two decimal places. Does not perform rounding.
rs_update_to_last_day	RULE SPECIFICATION	Updates an input date in the format yyyy-mm-dd to the final day of the month. Returns the date in the format yyyy-mm-dd hh:mm:ss.s.
rs_use_dict_to_replace_custom_string	RULE SPECIFICATION	Cross-references the input values with entries in a specified dictionary and replaces them with a custom string.
rs_verify_date1_precedes_date2	RULE SPECIFICATION	Verifies that date1 precedes date2.
v_Global_AddressValidation_Hybrid	VERIFIER	Validates the deliverability of address records from multiple countries and corrects errors where possible. The asset is configured to read input address fields in a hybrid structure.

CHAPTER 3

Data Quality bundle for financial services

Use Data Quality Financial Bundle to verify and enhance data from organizations in the financial services and mortgage loans sector. The Data Quality bundle for financial services includes rules that perform the following data quality operations: contact data cleansing, financial data cleansing, general data cleansing, matching and deduplication. The bundle also includes assets that evaluate mortgage data in the United States.

General assets

The following table lists the assets contained in the Data Quality bundle for financial services:

Name	Asset Type	Description
c_IFSC_Bank_Code	CLEANSE	Replaces an Indian Financial System Code (IFSC) with a valid full bank name. An IFSC is an eleven-character alphanumeric code that the Reserve Bank of India uses to identify bank branches within the NEFT (National Electronic Funds Transfer) network.
c_Merge_Currency_Country_Codes	CLEANSE	Merges the currency code and ISO three-character country code with an underscore character as a delimiter. For example, AFN_AFG.
c_Remove_Extra_Spaces	CLEANSE	Replaces multiple consecutive spaces with a single space and trims leading and trailing spaces.
c_Remove_Punctuation	CLEANSE	Removes all occurrences of punctuation from the input field.
c_Remove_Space	CLEANSE	Removes all occurrences of a character space from the input field.
c_Uppercase	CLEANSE	Converts the input text to uppercase.
c_US_Get_Name_Standard	CLEANSE	Standardizes given names.
dedupe_CompanyName_International_Match	DEDUPLICATE	Identifies duplicate records based on similarities in organization name data.

Name	Asset Type	Description
dedupe_Household_Name_Address_Match	DEDUPLICATE	Identifies duplicate records based on similarities in family name and address data. The operation uses Address Part1 as the index key.
dedupe_Individual_Name_and_Address	DEDUPLICATE	Identifies duplicate records based on similarities in person name and address data.
dedupe_Individual_Name_and_Email	DEDUPLICATE	Identifies duplicate records based on similarities in person name and email address data.
dedupe_Individual_Name_Date	DEDUPLICATE	Identifies duplicate records based on similarities in person name and date data.
dedupe_Individual_Name_Match	DEDUPLICATE	Identifies duplicate records based on similarities in person name data.
dedupe_Individual_Name_Phone	DEDUPLICATE	Identifies duplicate records based on person name and telephone number data.
dedupe_Match_Company_and_Address	DEDUPLICATE	Identifies duplicate records based on organization name and address data. The operation uses Company Name as the index key.
currency_codes_country_codes_infa	DICTIONARY	Contains a list of currency and country codes.
currency_codes_infa	DICTIONARY	Contains the currency codes of currencies around the world.
currency_symbol_from_code_infa	DICTIONARY	Contains a list of currency symbols and their corresponding currency names.
ccar_commercial_loan_flag	DICTIONARY	Contains commercial loan flag that is used in Comprehensive Capital Analysis and Review (CCAR) rule specifications.
ccar_false_flag	DICTIONARY	Contains false flag that is used in CCAR rule specifications.
ccar_N_flag	DICTIONARY	Contains 'No' flag that is used in CCAR rule specifications.
ccar_Occupancy_Code	DICTIONARY	Contains occupancy code that is used in CCAR rule specifications.
ccar_true_flag	DICTIONARY	Contains true flag that is used in CCAR rule specifications.
ccar_U_flag	DICTIONARY	Contains unknown flag that is used in CCAR rule specifications.
ccar_usa_state_three_digit_zip	DICTIONARY	Contains three-digit Zip Code and state code.
ccar_usa_zipcode_with_state	DICTIONARY	Contains a list of Zip Codes related to states.

Name	Asset Type	Description
ccar_Y_flag	DICTIONARY	Contains 'Yes' flag that is used in CCAR rule specifications.
Full_IFSC_Code_List	DICTIONARY	Contains a list of consolidated Indian Financial System Codes that identify NEFT-enabled bank branches as of 2022. NEFT abbreviates National Electronic Funds Transfer.
gbr_bank_sort_codes_infa	DICTIONARY	Contains United Kingdom bank sort codes.
iso_2chr_country_code_infa	DICTIONARY	Contains a list of two-character ISO country codes and the corresponding country names.
moodys_rating_codes_infa	DICTIONARY	Contains codes, grades, and description data that the Moody's rating agency uses.
s_and_p_rating_codes_infa	DICTIONARY	Contains Standard and Poors rating codes, grades, and description data.
usa_nicknames_infa	DICTIONARY	Contains a list of United States nicknames.
usa_states_infa	DICTIONARY	Contains a list of states in the United States and two-character codes.
user_defined_accounts_infa	DICTIONARY	Contains user-defined account data.
user_defined_ratings_infa	DICTIONARY	Contains user-defined rating codes, grades, and description data.
dq_Account_Validation	LABELER	Adds a label to an input row to indicate whether the contents of an account status field in the row is valid.
lbl_currency_codes	LABELER	Labels the valid currency codes as X.
lbl_currency_country_codes	LABELER	Labels a valid combination of currency code and country code. Returns Valid or Invalid.
lbl_IFSC_Bank_Code	LABELER	Labels Indian Financial System Code (IFSC) bank codes. If the input value is not valid, the asset returns the input value.
lbl_ifsc_code	LABELER	Labels Indian Financial System Code (IFSC) bank codes. Returns Valid or Invalid.
lbl_iso_Validate_Country_Code	LABELER	Labels ISO country codes.
lbl_Italian_Fiscal_Code	LABELER	Labels Italian fiscal codes.
lbl_Moodys_Rating	LABELER	Labels Moody's rating values.
lbl_Standard_and_Poors_Rating	LABELER	Labels Standard and Poor's rating values.
lbl_transit_number	LABELER	Assigns the label CAN_TRANSIT_NUM to values that match the Canadian transit number format.

Name	Asset Type	Description
lbl_User_Defined_Rating	LABELER	Labels user-defined rating values.
lbl_Validate_Country_Code	LABELER	Validates an ISO two-character country code. Returns Valid or Invalid.
lbl_Validate_CurrencyCode	LABELER	Validates an input currency code. Returns Valid or Invalid.
lbl_VAT_Italy	LABELER	Validates an Italian VAT (Value Added Tax) number format.
mplt_Account_Status_Validation	DMAPPLET	Validates account status data. The mapplet requires account status reference data.
mplt_Annual_Revenue_Validation	DMAPPLET	Checks that an input value is greater than zero and contains a currency symbol. Returns a status, description, the revenue value with currency and commas removed as a string, and the currency symbol if found at the start or end of the string.
mplt_Beta_Coefficient_Validation	DMAPPLET	Validates that a Beta coefficient value is a number. The mapplet indicates that the value is a positive number, negative number, zero, or not a number.
mplt_BIC_SWIFT_Code_Validation	DMAPPLET	Validates a Bank Identifier Code (BIC) or Society for Worldwide Interbank Financial Telecommunication (SWIFT) code by pattern recognition and country code validation.
mplt_CAN_Transit_Number_Validation	DMAPPLET	Validates the format of a Canadian transit number, as used for example in an EFT (electronic funds transfer) transaction.
mplt_Company_Name_and_Address_Match	DMAPPLET	Identifies duplicate records based on similarities in organization name and address data.
mplt_Company_Name_International_Match	DMAPPLET	Identifies duplicate records based on similarities in international organization name and address data.
mplt_Credit_Card_Security_Code_Validation	DMAPPLET	Validates that the credit card security code is a whole number that contains three or four digits.
mplt_Currency_Code_Country_Validation	DMAPPLET	Validates that the currency code is valid for the ISO three-character country code.
mplt_Currency_Code_Validation	DMAPPLET	Validates the currency code. Returns Valid or Invalid.
mplt_CUSIP_Validation	DMAPPLET	Validates CUSIP (Committee on Uniform Securities Identification Procedures) numbers. The mapplet determines the check digit for the input CUSIP number and compares it to valid CUSIP check digits.

Name	Asset Type	Description
mplt_Delta_Validation	DMAPPLET	Validates that the Delta value is positive, negative, or zero. The Delta is a ratio that compares the changes in the derivative price and the underlying asset price.
mplt_Dividend_Yield_Validation	DMAPPLET	Validates that the dividend yield value is a number greater than or equal to zero. The mapplet returns whether the value is a positive number, negative number, zero, or not a number.
mplt_EAD_Drawn_Balance_Validation	DMAPPLET	Validates that the amount listed as the exposure at default (EAD) is not less than the drawn balance. The mapplet follows the guidelines for EAD calculation from the Financial Services Authority in the United Kingdom.
mplt_EAD_Validation	DMAPPLET	Validates that the exposure at default (EAD) value is a number. The mapplet returns whether the value is a positive number, negative number, zero, or not a number.
mplt_EPS_Validation	DMAPPLET	Validates that the earnings per share (EPS) value is a number greater than or equal to zero.
mplt_Ex_Dividend_Date_Validation	DMAPPLET	Validates that the ex-dividend date and the record date are valid dates and that the ex-dividend date is earlier than the record date. The mapplet returns the difference in days between the record date and the ex-dividend date. The mapplet identifies dates with a difference of more than 15 days as not valid.
mplt_Familyname_and_Address_Match	DMAPPLET	Identifies duplicate rows based on surname and United States address data. The mapplet uses a combination of characters from the surname values and the postal code values to generate group keys.
mplt_Gamma_Validation	DMAPPLET	Validates that the Gamma string is a number. The mapplet returns whether the string is a positive number, negative number, zero, or not a number. Gamma measures the instability of the Delta value. A higher Gamma is an indication of a greater potential for change in the Delta value.
mplt_GBR_Bank_Account_Parse	DMAPPLET	Parses eight-digit numeric strings as United Kingdom bank account numbers.
mplt_GBR_Bank_Account_Validation	DMAPPLET	Validates United Kingdom bank account numbers. The mapplet returns codes that indicate whether the input is numeric and whether it contains the correct number of digits.
mplt_GBR_Bank_Sort_Code_Parse	DMAPPLET	Parses six-digit numeric strings as United Kingdom bank sort codes. The mapplet parses strings of numbers in the following formats: - Consecutive numbers (999999) - Numbers delimited with a dash (99-99-99)

Name	Asset Type	Description
mplt_GBR_Bank_Sort_Code_Standardize	DMAPPLET	Standardizes a United Kingdom bank sort code to the format NN-NN-NN.
mplt_GBR_Bank_Sort_Code_Validation	DMAPPLET	Validates the format and length of United Kingdom bank sort codes that are standardized to the dash-delimited format (99-99-99). The mapplet returns a Status field that describes the validity of the sort code and a Validation Note field that explains the status. If the sort code prefix matches a known assignment for a United Kingdom bank, the Validation Note field includes the bank name.
mplt_IND_Bank_IFSC_Code_Validation	DMAPPLET	Validates an Indian Financial System Code (IFSC) code.
mplt_IND_Bank_Name_From_IFSC_Code	DMAPPLET	Derives the bank name that an Indian Financial System Code (IFSC) represents.
mplt_Individual_Name_and_Address_Match	DMAPPLET	Identifies duplicate rows based on person names and United States address data. The deduplicate operation groups the data on surname and address values.
mplt_Individual_Name_and_Email_Match	DMAPPLET	Identifies duplicate rows based on person names and email addresses. The deduplicate operation groups the data on email address values.
mplt_Individual_Name_Date_Match	DMAPPLET	Identifies duplicate rows based on person names and date data. The deduplicate operation groups the data on date values.
mplt_Individual_Name_Match	DMAPPLET	Identifies duplicate rows based on person names. The deduplicate operation groups the data on surname values.
mplt_Individual_Name_Phone_Match	DMAPPLET	Identifies duplicate rows based on person names and telephone numbers. The deduplicate operation groups the data by telephone area code.
mplt_Interest_Rate_Within_Range	DMAPPLET	Validates if the decimal interest rate value is within the specified range. The range is set by two variable ports in an Expression transformation. The mapplet returns True or False.
mplt_ISIN_Code_Validation	DMAPPLET	Validates that the input value is an International Securities Identification Number (ISIN). The mapplet checks the structure of the value and verifies the check digit.
mplt_ITA_Codice_Fiscale_Validation	DMAPPLET	Validates an Italian fiscal code, or Codice Fiscale.
mplt_ITA_VAT_Number_Validation	DMAPPLET	Validates an Italian VAT (Value Added Tax) number.
mplt_Loan_to_Value_Ratio	DMAPPLET	Calculates the loan to value ratio, which is the loan amount divided by the property value.

Name	Asset Type	Description
mplt_Loss_Given_Default_Validation	DMAPPLET	Validates that the input is numeric and a positive, negative, or zero value. The mapplet is designed to assist in loss-given-default calculations for financial institutions.
mplt_Positive_Close_Price_Value_Validation	DMAPPLET	Validates that the input is a number greater than zero. The mapplet is designed to help financial institutions verify the closing prices of securities.
mplt_Positive_Coupon_Percent_Validation	DMAPPLET	Validates that the input is a number greater than zero. The mapplet is designed to help financial institution identify bond securities with positive interest rates.
mplt_Positive_Last_Price_Value_Validation	DMAPPLET	Validates that the input is a number greater than zero. The mapplet is designed to help financial institutions determine positive security valuations.
mplt_Positive_Number_Validation	DMAPPLET	Checks that the input is a number greater than zero.
mplt_Positive_Open_Price_Validation	DMAPPLET	Validates that the input is a number greater than zero. The mapplet is designed to help financial institutions verify the opening prices of securities.
mplt_Positive_Volume_Validation	DMAPPLET	Validates that the input is a number greater than zero. The mapplet is designed to help financial institutions identify positive changes in the volume of securities transactions.
mplt_Price_Earnings_Ratio_Validation	DMAPPLET	Validates that the price-to-earnings ratio is a positive number in the range of 0 - 100.
mplt_Probability_of_Default_Validation	DMAPPLET	Validates that the probability of default value is numeric and indicates if it is a positive, negative, or zero value. If positive, the mapplet returns status messages for values in the following ranges: - Less than or equal to 0.1. - Greater than 0.1 and less than or equal to 0.5. - Greater than 0.5 and less than or equal to 1. - Greater than 1.
mplt_Rating_Code_Validation	DMAPPLET	Validates that a rating is in the Standard & Poor's ratings scale, the Moody's ratings scale, or in a user-defined list.
mplt_Remove_Extra_Spaces	DMAPPLET	Replaces multiple consecutive character spaces with a single space, and trims leading and trailing spaces.
mplt_Remove_Punctuation	DMAPPLET	Removes all punctuation symbols from the input.
mplt_Remove_Punctuation_and_Space	DMAPPLET	Removes all punctuation symbols and character spaces from an input field, and returns the remaining digits and letters.
mplt_Remove_Space	DMAPPLET	Removes all occurrences of a character space from an input field.

Name	Asset Type	Description
mplt_UpperCase	DMAPPLET	Changes alphabetic characters to uppercase.
mplt_USA_Given_Name_Standard	DMAPPLET	Generates given names from United States nicknames. For example, the mapplet standardizes the nickname "Bob" to the given name "Robert."
mplt_Volatility_Validation	DMAPPLET	Validates that the volatility value is a number greater than or equal to zero.
rule_Convert_to_Number	DMAPPLET	Converts a string value to a number.
rule_Decimal_Length	DMAPPLET	Determines the length of the decimal portion of a number. Returns how many digits exist to the right of the decimal.
rule_Is_Date	DMAPPLET	Determines if an input string is a valid date. Returns Valid or Invalid.
rule_Is_Numeric	DMAPPLET	Determines if a string value consists of only numbers. Returns True or False.
p_GBR_Account_Parse	PARSE	Parses a United Kingdom bank account number from a string.
p_match_sort_codes	PARSE	Returns the name of the bank that uses the sort code in the input data.
p_Parse_BIC_Code	PARSE	Parses BIC or SWIFT codes from input data.
p_Parse_GBR_Bank_Sort_Code	PARSE	Parses six-digit strings as United Kingdom bank sort codes. The asset parses strings of numbers in the following formats: - Consecutive numbers (999999) - Numbers delimited with a dash (99-99-99)
p_Parse_GBR_Match_Bank_Sort_Codes	PARSE	Parses a United Kingdom bank sorting code from the input data when an input value matches a value in a dictionary.

United States Mortgage assets

The following table lists the assets that evaluate mortgage data in the United States:

Name	Asset Type	Description
ARM_Initial_Rate_datatype_limitation	RULE SPECIFICATION	Validates that the initial rate of an adjustable rate mortgage fits a decimal (7,5) datatype field or is an empty string.
ARM_Initial_Rate_Period_small_int	RULE SPECIFICATION	Validates that the initial rate period of an adjustable rate mortgage is a smallint value.
ARM_InitialRate_data_limit	RULE SPECIFICATION	Validates that the initial rate of an adjustable rate mortgage is not greater than or equal to 18 or less than 1.

Name	Asset Type	Description
ARM_InitRatePrd_not_poped_FxdRteLoans	RULE SPECIFICATION	Validates that the initial rate period of an adjustable rate mortgage is not populated for fixed-rate loans.
ARM_Lifetime_Rate_Cap_limitation	RULE SPECIFICATION	Validates that the lifetime rate cap on an adjustable rate mortgage fits a decimal (7,5) datatype field or is an empty string.
ARM_Lifetime_Rate_Floor_datatype_limit	RULE SPECIFICATION	Validates that the lifetime rate floor of an adjustable rate mortgage fits a decimal (7,5) datatype field or is an empty string.
ARM_loans_should_have_ARM_IRP	RULE SPECIFICATION	Validates that an adjustable rate mortgage loan has an initial rate period.
ARM_Margin_at_Orig_datatype_limit	RULE SPECIFICATION	Validates that the margin at origination of an adjustable rate mortgage fits a decimal (7,5) datatype field or is an empty string.
ARM_Margin_at_Origination_not_change	RULE SPECIFICATION	Validates that the margin at origination of an adjustable rate mortgage is not changed unless it is not provided.
ARM_Neg_Amortization_Limit	RULE SPECIFICATION	Validates that the negative amortization limit of an adjustable rate mortgage is not populated for fixed-rate loans.
ARM_Neg_Amortization_Limit_datatype_lim	RULE SPECIFICATION	Validates that the negative amortization limit of an adjustable rate mortgage fits a decimal (7,5) datatype field or is an empty string.
ARM_Neg_Amortization_Limit_empty	RULE SPECIFICATION	Validates that the negative amortization limit of an adjustable rate mortgage is an empty string if negative amortization is not allowed.
ARM_Negative_Amortization_Limit	RULE SPECIFICATION	Validates that the negative amortization limit of an adjustable rate mortgage is not greater than 125 percent.
ARM_Periodic_Pay_Cap_datatype_limit	RULE SPECIFICATION	Validates that the periodic pay cap on an adjustable rate mortgage fits a decimal (7,5) datatype field or is an empty string.
ARM_Periodic_Pay_Floor_datatype_limit	RULE SPECIFICATION	Validates that the periodic pay floor on an adjustable rate mortgage fits a decimal (7,5) datatype field or is an empty string.
ARM_Periodic_Rate_Cap_datatype_limit	RULE SPECIFICATION	Validates that the periodic rate cap on an adjustable rate mortgage fits a decimal (7,5) datatype field or is an empty string.
ARM_Periodic_Rate_Floor_datatype_limit	RULE SPECIFICATION	Validates that the periodic rate floor on an adjustable rate mortgage fits a decimal (7,5) datatype field or is an empty string.

Name	Asset Type	Description
Balance_LT_125_pct_LoanAmountDisp	RULE SPECIFICATION	Validates that the balance is less than 125 percent of the original loan amount disbursed.
Balloon_Term_limits	RULE SPECIFICATION	Validates that the balloon term is a number greater than or equal to 12 and less than or equal to 600 when the balloon flag is set.
Balloon_Term_must_be_a_smallint	RULE SPECIFICATION	Validates that the balloon term is a smallint value.
Balloon_Term_should_not_change	RULE SPECIFICATION	Validates that the balloon term is not changed unless it is not provided.
BalloonFlag_Is_Valid_Value	RULE SPECIFICATION	Validates that the balloon flag indicates the status Yes, No, or Unknown.
BalloonFlag_Unchanged	RULE SPECIFICATION	Validates that the balloon flag is not changed unless it indicates an unknown status or it is not provided.
Buydown_Flag_is_Unknown	RULE SPECIFICATION	Validates that the buydown flag indicates an unknown status.
Buydown_Flag_Must_Be_Valid_Value	RULE SPECIFICATION	Validates that the buydown flag indicates the status Yes, No, or Unknown.
Buydown_Flag_should_no_change	RULE SPECIFICATION	Validates that the buydown flag is not changed unless it indicates an unknown status or it is not provided.
Calc_LTV_GT_20pct_Change_from_Orig_LTV	RULE SPECIFICATION	Validates that the calculated loan-to-value ratio differs from the original loan-to-value ratio by more than 20 percent.
CBS_is_Number_Between_300_and_899	RULE SPECIFICATION	Validates that the origination credit bureau score is a valid whole number between 300 and 899.
CC_must_be_Specific_Value_or_empty_str	RULE SPECIFICATION	Validates that the credit class value is within a set of defined values.
Closing_Date_Check	RULE SPECIFICATION	Validates that the closing date is a valid date with the format `YYYY-MM-DD`. The date should not contain spaces and should not be null.
ClsngDate_Trm_3Mnth_Grtr_Than_Curr_Mnth	RULE SPECIFICATION	Validates that the closing date plus the loan term plus three months does not indicate a date that is earlier than the current month.
CLTV_Is_Number_GT_10pct_and_LT_150pct	RULE SPECIFICATION	Validates that the combined loan-to-value ratio is a number greater than 10 percent and less than 150 percent.

Name	Asset Type	Description
Credit_Class_Must_Be_Specific_Value	RULE SPECIFICATION	Validates that the credit class value is within a set of defined values.
Credit_class_should_not_change	RULE SPECIFICATION	Validates that the current credit class value matches the credit class value from the previous month.
FirstPayDate_Before_NextDueDate	RULE SPECIFICATION	Validates that the first payment date is not later than the next payment date.
FirstPayDate_date_restrictions	RULE SPECIFICATION	Validates that the first payment date is not more than three months later than the current month or more than 40 years ago.
FirstPayDate_Valid_Date	RULE SPECIFICATION	Validates that the input string is a valid date value in the format `YYYYMMDD`.
Fixed_loans_not_have_an_ARM_IRAP	RULE SPECIFICATION	Validates that a fixed loan does not have an adjustable rate mortgage initial rate adjustment period.
Fixed_Loans_not_have_ARM_IRP	RULE SPECIFICATION	Validates that a fixed loan does not have an adjustable rate mortgage initial rate period.
For_NCL_OPV_must_be_valid_positive_num	RULE SPECIFICATION	For a non-commercial loan, validates that the original property value is a valid positive whole number that is greater than or equal to 1,000 and less than or equal to 10,000,000.
FPD_LessThanOrEqual_ClosingDate	RULE SPECIFICATION	Validates that the closing date is earlier than or equal to the first payment date.
Frontend_DTI_is_Positive_Whole_Number	RULE SPECIFICATION	Validates that the original front-end debt to income value is a valid positive whole number.
If_not_a_Balloon_not_have_a_Balloon_Trm	RULE SPECIFICATION	Validates that a loan that is not a balloon loan does not have a balloon term.
Interest_Only_At_Origin_cannot_change	RULE SPECIFICATION	Validates that the interest-only mortgage at origination value is not changed unless it is unknown or it is not provided.
Interest_Only_At_Origination_is_Unknown	RULE SPECIFICATION	Validates that the interest-only mortgage at origination value is unknown.
Lien_Position_At_Origin_Must_Be_1	RULE SPECIFICATION	Validates that the lien position at origination is 1.
Lien_Position_at_Origin_No_change	RULE SPECIFICATION	Validates that the lien position at origination is not changed unless it is unknown or it is not provided.
Lien_Position_at_Origin_Specific_Value	RULE SPECIFICATION	Validates that the lien position at origination is within a set of defined values.

Name	Asset Type	Description
Loan_Closing_Date_Less_or_Equal_40_Yrs	RULE SPECIFICATION	Validates that the year of the loan closing date is not more than 40 years ago.
Loan_Closing_Month_Not_Greater_Curr_Mnt	RULE SPECIFICATION	Validates that the month of the loan closing date is earlier than the current month.
Loan_Number_Length_Check	RULE SPECIFICATION	Validates that the loan number is not empty or null and does not contain more than 32 characters.
Loan_Type_is_Other	RULE SPECIFICATION	Validates that the loan is a commercial loan.
Loan_Type_is_Unknown_or_Other	RULE SPECIFICATION	Validates that the loan is a commercial loan or the loan type is unknown.
LoanAmount_Unchanged_if_NotEmpty	RULE SPECIFICATION	Validates that the original loan amount disbursed is not changed unless the amount is not provided.
LoanPurpose_is_Other_or_Unknown	RULE SPECIFICATION	Validates that the loan purpose is not defined or is unknown.
LoanPurpose_Unchanged	RULE SPECIFICATION	Validates that the loan purpose coding value is not changed unless it is unknown or it is not provided.
Loans_with_ARM_Init_Rate_Period_is_ARM	RULE SPECIFICATION	Validates that a loan with an adjustable rate mortgage initial rate period is an adjustable rate mortgage.
LoanSource_is_Unknown	RULE SPECIFICATION	Validates that the loan source is unknown.
LoanSource_Unchanged	RULE SPECIFICATION	Validates that the loan source is not changed unless it is unknown or it is not provided.
LoanType_Unchanged	RULE SPECIFICATION	Validates that the loan type is not changed unless it is unknown or it is not provided.
LVT_GT_prcnt_range_for_NonHARP_and_HARP	RULE SPECIFICATION	Validates that the original loan-to-value ratio is a number greater than ten percent and less than 150 percent for HARP (Home Affordable Refinance Program) loans or a number greater than ten percent and less than 125 percent for non-HARP loans.
MI_Coverage_Pct_At_Origin_no_change	RULE SPECIFICATION	Validates that mortgage insurance cover percentage at origination is not changed unless the percentage is not provided.
MI_Coverage_Percent_At_Origin_datatype	RULE SPECIFICATION	Validates that mortgage insurance cover percentage at origination fits a decimal (4, 2) datatype field or is an empty string.

Name	Asset Type	Description
MI_Coverage_Prcnt_Origin_btwn_1_50_pct	RULE SPECIFICATION	Validates that mortgage insurance cover percentage at origination is a value between one percent and 50 percent.
Neg_Amortization_Flag_be_Specifc_Value	RULE SPECIFICATION	Validates that the negative amortization flag value is within a set of defined values or is an empty string.
Neg_Amortization_Flag_no_change	RULE SPECIFICATION	Validates that the negative amortization flag value is not changed unless it is not provided.
Negative_Amortization_Flag_be_Y_or_N	RULE SPECIFICATION	Validates that the negative amortization flag value indicates yes or no.
NonHARP_Loan_and_LTV_GTE_125pct	RULE SPECIFICATION	Validates that a loan is not a HARP loan and the calculated original loan-to-value ratio is greater than or equal to 125 percent.
Num_Of_Units_not_change	RULE SPECIFICATION	Validates that the number of units is not changed unless it is unknown or it is not provided.
Number_of_Units_is_Unknown_or_Other	RULE SPECIFICATION	Validates that the loan is a commercial loan and the number of units is not specified.
NumOfUnits_is_Specific_Value_or_Unknow	RULE SPECIFICATION	Validates that the number of units is within a set of defined values.
Occupancy_Code_Must_be_Specific_Value	RULE SPECIFICATION	Validates that the occupancy code is within a set of defined values
Occupancy_is_Unknown	RULE SPECIFICATION	Validates that the occupancy value is unknown.
Occupancy_no_change	RULE SPECIFICATION	Validates that the occupancy value is not changed unless it is unknown or it is not provided.
Opt_ARM_At_Origin_same_as_prior_month	RULE SPECIFICATION	Validates that the option adjustable rate mortgage at origination flag is same as the previous month.
Opt_ARM_Origin_Spec_Valu_or_an_emty_str	RULE SPECIFICATION	Validates that the option adjustable rate mortgage at origination value is Y or N or an empty string.
Option_ARM_Origin_must_be_Spec_Value	RULE SPECIFICATION	Validates that the option adjustable rate mortgage at origination value is Y or N.
Orig_Backend_DTI_Unchanged_If_Provided	RULE SPECIFICATION	Validates that the original back-end debt-to-income ratio is not changed unless it is not provided.

Name	Asset Type	Description
Orig_CBS_no_change	RULE SPECIFICATION	Validates that the origination credit bureau score is not changed unless it is not provided.
Orig_Frontend_DTI_Unchanged_If_Provided	RULE SPECIFICATION	Validates that the original Frontend DTI is not changed unless it is not provided
Orig_Loan_Term_number_between_12_600	RULE SPECIFICATION	Validates that the original loan term is a number between 12 and 600.
Orig_LoanAmount_Disbursed_be_integer	RULE SPECIFICATION	Validates that the original loan amount disbursed is an integer value.
Orig_LoanAmount_Disbursed_valid_pos_num	RULE SPECIFICATION	Validates that the original loan amount disbursed is a valid positive whole number.
Original_Backend_DTI_is_Positive_Number	RULE SPECIFICATION	Validates that the original back-end debt-to-income ratio is a valid positive whole number and is less than 100.
Original_Backend_DTI_is_SmallInt	RULE SPECIFICATION	Validates that the original back-end debt-to-income ratio is a smallint value.
Original_CLTV_Is_4_2_format_or_Is_Empty	RULE SPECIFICATION	Validates that the original combined loan-to-value ratio fits a decimal (4, 2) datatype field or is an empty string.
Original_CLTV_Not_Provided	RULE SPECIFICATION	Validates that the original combined loan-to-value ratio is not provided.
Original_CLTV_Unchanged_If_Not_Empty	RULE SPECIFICATION	Validates that the original combined loan-to-value ratio is not changed unless it is not provided.
Original_Frontend_DTI_GT_Backend_DTI	RULE SPECIFICATION	Validates that the original front-end debt-to-income ratio is greater than the back-end debt-to-income ratio.
Original_Frontend_DTI_is_Smallint	RULE SPECIFICATION	Validates that the original front-end debt-to-income ratio is a smallint value.
Original_LVT_Is_Number_with_4_2_format	RULE SPECIFICATION	Validates that the original loan-to-value ratio fits a decimal (4, 2) datatype field or is an empty string.
Original_LVT_Unchanged_If_Not_Empty	RULE SPECIFICATION	Validates that the original loan-to-value ratio is not changed unless it is not provided.
Original_Property_Value_an_Integer	RULE SPECIFICATION	Validates that the original property value is an integer.
OriginalPropertyValue_Unchanged	RULE SPECIFICATION	Validates that the original property value is not changed unless it is not provided.
Origination_CBS_a_smallint	RULE SPECIFICATION	Validates that the origination credit bureau score is a smallint value.

Name	Asset Type	Description
Origination_CBS_between_300_899	RULE SPECIFICATION	Validates that the origination credit bureau score is a value between 300 and 899.
Origination_CBS_should_not_change	RULE SPECIFICATION	Validates that the origination credit bureau score is not changed unless it is not provided.
Prop_ZIPCode_Less_than_equal_to_5_chars	RULE SPECIFICATION	Validates that the property Zip Code contains five or more characters or is an empty string.
Property_type_cannot_change	RULE SPECIFICATION	Validates that the property type is not changed unless it is unknown or it is not provided.
PropertyState_Length_Not_GreaterThan_2	RULE SPECIFICATION	Validates that the state code for the property contains not more than two characters.
PropertyType_is_Unknown	RULE SPECIFICATION	Validates that the property type is unknown.
Recourse_Flag_is_Unknown	RULE SPECIFICATION	Validates that the recourse flag indicates an unknown status.
RecourseFlag_Unchanged	RULE SPECIFICATION	Validates that the recourse flag is not changed unless it is unknown or it is not provided.
State_Abbr_Valid_for_NonCommercial_Loan	RULE SPECIFICATION	For non-commercial loans, validates that the property state is a valid postal abbreviation of one of the 50 states, or DC, or a valid United States possession.
ThreeDigit_ZipCode_Consisent_with_State	RULE SPECIFICATION	Validates that the first three digits of a Zip Code are consistent with a state in the United States.
Zipcode_Unchanged_and_NotEmpty	RULE SPECIFICATION	Validates that the property Zip Code is not changed unless it is not provided.
ZipCode5_for_NonCommercial_Loan	RULE SPECIFICATION	For non-commercial loans, validates that the property Zip Code contains five digits.

CHAPTER 4

Data Quality bundle for crisis responses

The Crisis Response accelerator can help you find and fix errors and inconsistencies in your data, parse meaningful information from your data, improve your data structure, and quickly build a data quality project to analyze and manage your data in a crisis such as COVID-19.

Cloud Data Quality Crisis Response Bundle

You can deliver a coordinated, data-driven response in your organization.
The following table describes the crisis response bundle assets:

Name	Type	Description
blood_type_infa	Dictionary	Contains blood groups and their alternative names.
c_av6_address_type_desc	Cleanse	Returns descriptions for the Address Type Codes.
c_av6_result_quality_desc	Cleanse	Returns descriptions for Result Quality.
c_av6_verification_status_code_desc	Cleanse	Returns descriptions for verification status codes.
c_blood_type_standardization	Cleanse	Standardizes blood type values.
c_country_id_dq_Prep_for_Country_Parse	Cleanse	Removes the short forms of street values. For example, removes Rd, St, Blvd.
c_Country_ISO2Char	Cleanse	Replaces country names with their respective two-letter ISO country codes.
c_Country_ISO3Char	Cleanse	Replaces country names with their respective three-letter ISO country codes.
c_cpt_code_from_section	Cleanse	Replaces valid Current Procedural Terminology (CPT) descriptions with their respective codes.

Name	Type	Description
c_dq_Label_US_City_Postcode	Cleanse	Replaces city and ZIP Code values of the United States with dictionary values.
c_dq_Label_US_State_Postcode	Cleanse	Labels United States postcodes.
c_dq_State_Validate	Cleanse	Validates if the input is a valid state name in the United States.
c_dq_Validate_County	Cleanse	Validates if the input is a valid county name in the United States.
c_icd10_Codes_mplt_Association	Cleanse	Replaces ICD-10 code descriptions with their respective codes. ICD-10 identifies the tenth revision of the International Statistical Classification of Diseases and Related Health Problems.
c_icd10_Desc_mplt_Association	Cleanse	Replaces ICD-10 codes with their respective descriptions.
c_iso_full_country_name_infa	Cleanse	Replaces country identifiers in the input string with their respective full ISO country names.
c_Label_Default_Date_Patterns	Cleanse	Replaces dictionary matches for default date patterns with the term 'Valid'.
c_lbl_US_State_Locality	Cleanse	Replaces a valid United States state name and county name combination with the label 'valid_usa'.
c_Out_Country_Name	Cleanse	Standardizes country names.
c_predirectional_std	Cleanse	Standardizes common descriptive terms in address data, for example sub-building and directional indicators.
c_Remove_Apostrophe	Cleanse	Removes an apostrophe from the input.
c_Remove_Extra_Spaces	Cleanse	Replaces multiple consecutive spaces with a single space, and trims leading and trailing spaces.
c_Remove_Hyphen	Cleanse	Removes any hyphen from a string.
c_remove_leading_zero	Cleanse	Removes any leading zero from a string.
c_Remove_Parentheses	Cleanse	Removes all occurrences of left and right parenthesis from a string.
c_Remove_Period	Cleanse	Removes a period from a string.
c_Remove_Period_Parentheses	Cleanse	Removes all occurrences of a period and left and right parentheses from a string.

Name	Type	Description
c_remove_punctuation	Cleanse	Removes all punctuation from the input.
c_Remove_Quotation	Cleanse	Removes all double quotation marks from a string.
c_remove_single_quotes	Cleanse	Removes single quotation marks from a string.
c_Remove_Slashes	Cleanse	Removes all occurrences of forward and backward slashes from a string.
c_remove_space	Cleanse	Removes all occurrences of a space from a string.
c_replace_hypen_w_single_quotes_spc	Cleanse	Replaces hyphens with single quotation marks.
c_replace_mult_spaces_w_single_space	Cleanse	Replaces instances of two or more consecutive character spaces in an input field with a single space, and trims the leading and trailing spaces.
c_Rmv_Dic_Matches_InstitutionMatch	Cleanse	Removes institution names that match the values in a dictionary.
c_standardize_rplce_bck_slash_with_fwd	Cleanse	Replaces a forward slash with a backward slash.
c_std_blood_mplt_Association	Cleanse	Standardizes blood type values.
c_Std_icd10_Block_Desc	Cleanse	Returns the ICD-10 block for a given medical condition based on its ICD-10 code.
c_std_icd10_Title_Description	Cleanse	Returns the ICD-10 title for a given medical condition based on its ICD-10 code.
c_std_terms_mplt_association	Cleanse	Replaces phrases or words associated with COVID-19 with the phrase 'COVID-19 virus identified'.
c_titlecase	Cleanse	Converts the input text to title case.
c_uppercase	Cleanse	Converts the input text to upper case.
c_US_Get_Name_Standard	Cleanse	Standardizes first names.
c_USA_PostCodes	Cleanse	Validates if the input has a valid ZIP Code against a reference table of United States ZIP Codes.
c_validate_area_code_label	Cleanse	Validates area codes.
c_validation_domain	Cleanse	Validates the domain of an email address.

Name	Type	Description
can_localities	Dictionary	Contains a list of Canadian localities.
can_province_and_postcode	Dictionary	Contains a list of Canadian provinces and postcodes.
can_provinces_infa	Dictionary	Contains a list of Canadian provinces.
comorbidity_space_replace_infa	Dictionary	Replaces a character space in comorbidity data with an underscore character.
cpt_codes_infa	Dictionary	Contains a list of CPT (Current Procedural Terminology) codes and their corresponding tests, surgeries, evaluations, or medical procedures.
cpt_section_by_code	Dictionary	Contains a list of CPT (Current Procedural Terminology) codes and the medical categories to which they belong, such as anesthesiology, surgery, radiology, pathology, and laboratory medicine.
diacritics_translate_infa	Dictionary	Contains a list of diacritic symbols.
domain_names_infa	Dictionary	Contains a list of domain names.
dq_av_mailabilityscores_infa	Dictionary	Contains Mailability Score address verification values.
dq_av_match_code_descriptions_infa	Dictionary	Contains Match Code address verification values and the list of process modes that return the values.
dq_av6_result_quality_descriptions	Dictionary	Contains a list of the result quality codes that Address Verification version 6 engine returns with descriptions.
dq_av6_status_code_descriptions_infa	Dictionary	Contains alpha status values that indicate the outcome of the verification operation for an address. Each status value has a corresponding text description.
facility_noise	Dictionary	Contains a list of healthcare facilities.
household_infa	Dictionary	Contains a list of household units.
icd10_adult_codes_infa	Dictionary	Contains a list of ICD-10-CM (Clinical Modification) codes that describe conditions for patients aged 15 and over.
icd10_as_principal_diagnosis_codes	Dictionary	Contains a list of ICD-10 codes and their descriptions.
icd10_cm_pcs_codes_check_infa	Dictionary	Contains a list of ICD-10- Procedure Coding System codes.

Name	Type	Description
icd10cm_codes_2020_classification_infa	Dictionary	Contains ICD-10 codes with their classifications.
icd10cm_codes_2020_infa	Dictionary	Contains a list of ICD-10-CM codes.
iso_full_country_name_infa	Dictionary	Contains ISO country names and their respective two-letter and three-letter ISO codes.
mplt_Association_btw_Infection_and_Blood_type	Mapplet	Standardizes blood type and diagnosis information and returns the corresponding ICD-10 codes and descriptions.
mplt_AV6_Assign_DQ_Status_Code_Description	Mapplet	Describes the Match Score value that a verifier asset returns for the Address Verification version 6 engine.
mplt_Blood_Type_Standardization	Mapplet	Standardizes input blood type values.
mplt_comparison_Dates	Mapplet	Compares two dates, calculates the difference between them, and generates four outputs. Each output represents the total difference between the date values in hours, days, months, or years. The output values are exclusive of each other are should not be used together to represent the difference between the data values. Input data type must be a string.
mplt_completeness	Mapplet	Checks if the input field is empty or null.
mplt_Completeness_Multi_Port	Mapplet	Checks if the input field is empty or null. Returns the incomplete data. Contains ten input and output ports.
mplt_Country_Name_Standardization1	Mapplet	Standardizes an input country name. Returns the full ISO country name and the corresponding ISO three-character and two-character codes.
mplt_CPT_code_Parse	Mapplet	Parses the current procedural terminology (CPT) code for all categories from a string.
mplt_CPT_code_validate	Mapplet	Validates a current procedural terminology (CPT) code, and returns the respective section codes.
mplt_Create_Institution_Match_Field	Mapplet	Standardizes institution names for use in a rule that performs match analysis.
mplt_Date_Complete	Mapplet	Checks if an input date string matches any user-defined default values. Does not validate the input string as a date.

Name	Type	Description
mplt_Date_Of_Birth_Validation	Mapplet	Checks the number of years between the input date and the current date. Returns whether the years represent an adult or a minor. Returns Valid for 0 - 120 inclusive and returns Invalid for all other entries.
mplt_Date_Parse	Mapplet	Parses a date from a string of text. Dates can be in dd/mm/yyyy, mm/dd/yyyy, or yyyy/dd/mm formats. Dashes, spaces, forward slashes, and dots can be used as separators in the date fields.
mplt_Date_Standardization	Mapplet	Formats an input date string using user-defined variables. Includes the exp_ParseDate and exp_FormatDate transformations. Input dates that are invalid or do not fit the designated input format are written as all zeros.
mplt_Date_Validation	Mapplet	Validates a given date.
mplt_Date_Validation_Variable_Format	Mapplet	Validates an input date string. Allows the user to specify the input date format.
mplt_Days_Between_Dates	Mapplet	Calculates the difference between the two dates in days. The result is the difference between the end date and the start date. Set the input datatype as String.
mplt_Days_From_Current_Date	Mapplet	Calculates the number of days between an input date and the current date (system date). Set the expected input type as String.
mplt_Email_Parse	Mapplet	Parses an email from within a string of text.
mplt_Email_validation	Mapplet	Validates the email format. Does not check if the email address is active. Returns Valid or Invalid.
mplt_Global_Address_Validation_Hybrid	Mapplet	Validates the deliverability of address records from multiple countries. The mapplet corrects errors where possible. Use the rule with input address fields in a hybrid format.
mplt_icd10_Adult_Code_Check	Mapplet	Checks if an ICD-10 code is for an adult patient, that is, for a patient in the age range of 15 through 124 inclusive. Returns 'True' if it is an adult code and 'False' if not.
mplt_icd10_Code_Definition	Mapplet	Returns the title and block definitions for an ICD-10 code.

Name	Type	Description
mplt_icd10_Code_Validation	Mapplet	Validates an ICD-10 code through a list that the user builds.
mplt_icd10_Questionable_Principal_Diagnosis_Code_Check	Mapplet	Checks if an ICD-10 code is determined to be insufficient for admission to an acute care hospital when used as a principal diagnosis. Returns 'True' if it is a Questionable Principal Diagnosis Code, 'False' if not.
mplt_Identify_Suspect_Names	Mapplet	Flags suspect or fake names.
mplt_IsNumeric	Mapplet	Verifies that the input data is numeric. The rule returns "True" or "False."
mplt_LowerCase	Mapplet	Converts the case of the input text to lowercase.
mplt_Match_Diag_Single	Mapplet	Compares morbidity values in two input strings, and returns a match score that represents the degree of similarity between them.
mplt_Negative_Number_Validation	Mapplet	Validates if the input is a negative number.
mplt_Non_Alphabetic_with_Space	Mapplet	Replaces all characters that are numbers or punctuations with a single space. Trims all consecutive spaces to a single space.
mplt_Parse_First_Word	Mapplet	Parses the first word in an input string to a field that the rule specifies.
mplt_Parse_Number_At_Start_Of_Line	Mapplet	Parses any number that occurs at the start of an input string to a field that the rule specifies. The rule reads strings from left to right.
mplt_Parse_Text_Between_Parentheses	Mapplet	Parses strings that are enclosed in parentheses to a field that the rule specifies. The rule contains an output field for the parsed strings and an output field for the input text without the parsed strings.
mplt_Parse_Text_in_Single_Quotes	Mapplet	Parses strings that are enclosed in quotation marks to a field that the rule specifies. When the input data contains multiple quoted elements, the rule parses the final element. The rule reads the input strings from left to right. The rule contains an output field for the parsed strings and an output field for the input text without the parsed strings.
mplt_Past_Date_Label	Mapplet	Determines whether an input date is earlier than the system date or later than the system date.

Name	Type	Description
mplt_Positive_Number_Validation	Mapplet	Verifies that the input data is a positive number.
mplt_Prename_Assingment	Mapplet	Assigns a prename based on the gender. Uses a variable to change a female prename assignment value. For example, Ms or Mrs.
mplt_Prepend_Zero_to_Single_digit	Mapplet	Prepends the numeral "0" to a single numeric character. Returns the original input if the input contains other characters.
mplt_Remove_All_Leading_Zeros	Mapplet	Removes all leading zeros from a string.
mplt_Remove_Apostrophe	Mapplet	Removes apostrophes. The rule merges the text strings on either side of the apostrophe.
mplt_Remove_Control_Characters	Mapplet	Removes control characters from text strings. The rule returns a string that contains the control characters and a string that contains the input text without the control characters.
mplt_Remove_Extra_Spaces	Mapplet	Replaces multiple consecutive spaces with a single space, and trims leading and trailing spaces.
mplt_Remove_Hyphen	Mapplet	Removes hyphens from anywhere in the input string.
mplt_Remove_Leading_Zero	Mapplet	Removes a leading zero from a string. Only removes the zero, and ignores all other occurrences.
mplt_Remove_Limited_Punctuation	Mapplet	Removes extraneous characters. Extraneous characters include slashes, back slashes, periods, exclamation marks, and underscores. The rule also replaces multiple consecutive spaces with a single space.
mplt_Remove_Non_Numbers	Mapplet	Removes characters that are neither a digit nor a number from the input string.
mplt_Remove_Parentheses	Mapplet	Removes all occurrences of left and right parentheses from a string.
mplt_Remove_Period	Mapplet	Removes all occurrences of a period from a string.
mplt_Remove_Period_Parantheses	Mapplet	Removes all occurrences of a period and left and right parentheses found in a string.

Name	Type	Description
mplt_Remove_Punctuation	Mapplet	Removes all punctuations from the input.
mplt_Remove_Punctuation_and_Space	Mapplet	Removes all punctuations and spaces.
mplt_Remove_Quotation	Mapplet	Removes all double quotation marks from a string. Does not replace with spaces.
mplt_Remove_Slashes	Mapplet	Removes all occurrences of forward and backward slashes in a string.
mplt_Remove_Space	Mapplet	Removes all occurrences of a space from an input field.
mplt_Replace_Limited_Punct_with_Space	Mapplet	Replaces the following punctuation characters with a single space: dash, back slash, period, exclamation mark, and underscore The rule also replaces two, three, and four consecutive spaces with a single space.
mplt_Salutation_Assignment	Mapplet	Generates formal and casual greetings from prenames and name tokens. For example, when input data contains "Mr. John Smith," the rule generates the formal greeting "Dear Mr. Smith," and the casual greeting "Dear John,". You can change the prefix and punctuation by editing the variables in the dq_Generate_Salutation Expression transformation.
mplt_String_Completeness	Mapplet	Checks a string for completeness. The rule also searches the input strings for values in the reference tablestring_default_values_infa. The reference table contains values such as NA, DEFAULT, and XX. If an input string contains a value in the reference table, the rule identifies the string as incomplete.
mplt_TitleCase	Mapplet	Converts the case of the input text to title case.
mplt_UpperCase	Mapplet	Returns all alphabetic characters in upper case.
mplt_UpperCase1000	Mapplet	Returns alphabetic characters in upper case. The input and output fields in the rule use a precision of 1,000.
mplt_USA_County_Validation	Mapplet	Checks if input string is a valid county name in the United States.

Name	Type	Description
mplt_USA_Gender_Assingment	Mapplet	Assigns gender according to first name. The mapplet returns "M" for male names, "F" for female names, and "U" if the gender is unknown. For example, the mapplet assigns the name "John Smith" a gender of "M" for male.
mplt_USA_Given_Name_Standard	Mapplet	Returns standard names for common nicknames. For example, returns Robert for Bob.
mplt_USA_Patient_Gender_Assignment	Mapplet	Assigns gender to a record according to the first-name information that the record contains. The mapplet returns "M" for male names, "F" for female names, and "U" if the gender is unknown. For example, the rule assigns the name "John Smith" a gender of "M" for male.
mplt_USA_Phone_Number_Parse	Mapplet	Parses a phone number from a text string in to its own field. The string must follow the standard format for the United States.
mplt_USA_Phone_Number_Standardization	Mapplet	Standardizes United States telephone numbers. The mapplet returns the telephone number in the following formats: - Standard - (nnn) nnn-nnnn - Dashes - nnn-nnn-nnnn - No spaces - nnnnnnnnnn - Phone_w_Extension - (nnn) nnn-nnnn ext nnnnn
mplt_USA_Phone_Number_Validation	Mapplet	Validates the area code and length of United States telephone numbers. The mapplet returns the following values: - 0 - NULL or blank input - 10 - Invalid length - 40 - Valid length, invalid area code - 95 - Valid length and area code
mplt_USA_Phone_Parse_Standardize_Validate	Mapplet	Parses a United States phone number from a text string, validates the area code, and returns the valid phone numbers in three different formats. Also returns a status value (valid or invalid) and a status description.
mplt_USA_Phone_w_Extension_Parse	Mapplet	Parses a telephone number, including an extension, that follows the standard number format for the United States.

Name	Type	Description
mplt_USA_Provider_Gender_Assignment	Mapplet	Assigns gender to a record according to the first-name information that the record contains. The mapplet returns "M" for male names, "F" for female names, and "U" if the gender is unknown. For example, the rule assigns the name "John Smith" a gender of "M" for male.
mplt_USA_SSN_Parse	Mapplet	Parses a United States Social Security number (SSN) from a string of text. The mapplet can parse numbers with or without dashes.
mplt_USA_SSN_Standardization	Mapplet	Standardizes United States Social Security numbers. The mapplet can output the number with no punctuation, with character spaces, and with dashes. Change the variable SSN_Format in the dq_SSN_Format transformation as desired. The default output is without punctuation.
mplt_USA_State_Validation	Mapplet	Validates the input string as a United States state name.
mplt_USA_ZIPCode_Validation	Mapplet	Validates that the input string is a five-digit United States ZIP Code.
mplt_Years_Since_Date_Of_Birth	Mapplet	Calculates the number of years since the input date.
name_apostrophe_infa	Dictionary	Contains apostrophe for names.
name_comma_infa	Dictionary	Contains commas for names.
name_dash_infa	Dictionary	Contains dashes for names.
p_AV_MailabilityScores_infa	Parse	Parses a mailability score value that appears in the DQ_AV_MailabilityScores_infa dictionary.
p_Cpt_Code_Parse	Parse	Parses a valid CPT code.
p_Date_Parse	Parse	Parses a date.
p_email_from_text	Parse	Uses a regular expression to parse an email address from text.
p_First_Word	Parse	Parses the first word in a string.
p_icd10_Adult_Code	Parse	Parses an ICD-10 adult code and description.
p_icd10_code_check_validate	Parse	Parses an ICD-10 code.

Name	Type	Description
p_icd10_questionable_diagnosis_parse	Parse	Parses an ICD-10 questionable diagnosis code or description.
p_icd10cm_Code	Parse	Parses an ICD-10-CM code from.
p_icd10cm_Descr	Parse	Parses an ICD-10-CM description from.
p_lbl_cpt_code_validate	Parse	Parses a CPT code.
p_Number_At_Start_of_String	Parse	Parses a number from the start of a string.
p_Parse_Country	Parse	Parses a three-character ISO country code.
p_suspect_names	Parse	Uses a dictionary to parse suspect names.
p_Text_btw_Parentheses	Parse	Parses the text between parentheses.
p_Text_btw_Single_Quotes	Parse	Parses the text between single quotes.
p_usa_gender_assignemnt_FN	Parse	Assigns gender based on the first name.
p_usa_gender_assingment_MN	Parse	Assigns gender based on the middle name.
p_USA_gender_Assingment_Prename	Parse	Assigns gender based on the prename.
p_USA_Phone	Parse	Parses a United States telephone number.
p_usa_phone_w_extension	Parse	Parses a United States telephone number with an extension from a larger string.
p_USA_SSN	Parse	Parses a United States Social Security Number in the format nnn-nn-nnnn.
rs_Assign_DQ_Match_Code_Description	Rule Specification	Assigns a description to a Match Code output from address verification.
rs_Cleanse_icd10cm_Label	Rule Specification	Provides a label for valid values in the ICD-10CM dictionary.
rs_Cpt_Validate	Rule Specification	Labels the input as a valid or invalid CPT code.
rs_dq_label_suspect_names	Rule Specification	Validates that the input is not a suspect name. Returns Suspect if a suspect name is found. Otherwise, returns Null.
rs_dq_output_match_code_desc	Rule Specification	Assigns a description to a Match Code output from address verification.

Name	Type	Description
rs_icd10_code_valid	Rule Specification	Checks for null values.
rs_icd10_Questionable_Diagnosis_Label	Rule Specification	Validates that the input contains an ICD-10 questionable diagnosis label.
rs_icd10cm_token1	Rule Specification	Provides the dictionary values for valid inputs based on the ICD-10CM dictionary.
rs_iso_full_country_name	Rule Specification	Returns the ISO country name for an ISO country abbreviation in the input data.
rs_Null_Check	Rule Specification	Checks for nulls and returns 'True' or 'False'.
rs_remove_multiple_spaces	Rule Specification	Replaces instances of two or more consecutive character spaces in an input field with a single space.
rs_uppercase	Rule Specification	Converts strings to uppercase.
suspect_names_infa	Dictionary	Contains a list of suspect names.
title_case_excptn_infa	Dictionary	Contains a list of strings that a rule can treat as exceptions in a case conversion operation.
usa_adrs_prfx_std_s_infa	Dictionary	Standardizes address name prefix in the United States.
usa_business_word_infa	Dictionary	Contains a list of business terms in use in the United States.
usa_company_names_std_infa	Dictionary	Contains a list of United States organization names.
usa_company_sufx_abrv_infa	Dictionary	Contains a list of corporate suffix abbreviations.
usa_connectors_infa	Dictionary	Contains a list of connectors used in the United States.
usa_counties_infa	Dictionary	Contains a list of United States county names. Includes any alternative names in use for the counties.
usa_firstnames_infa	Dictionary	Contains a list of first names in the United States.

Name	Type	Description
usa_gender_infa	Dictionary	Contains a list of United States first names and their commonly-assigned genders.
usa_locality_and_postcode	Dictionary	Contains a list of United States cities and towns and their associated ZIP Codes.
usa_name_designator_infa	Dictionary	Contains a list of name designators in use in the United States.
usa_name_frequencies_infa	Dictionary	Contains a list of United States names along with their frequencies.
usa_name_prefix_infa	Dictionary	Contains a list of name prefixes in use in the United States.
usa_name_suffix_infa	Dictionary	Contains a list of name suffixes in use in the United States.
usa_nicknames_infa	Dictionary	Contains a list of nicknames in use in the United States.
usa_occupation_infa	Dictionary	Contains a list of common occupations in the United States.
usa_postcodes	Dictionary	Contains a list of United States ZIP Codes.
usa_prename_gender_infa	Dictionary	Contains a list of United States first names and their commonly-assigned genders.
usa_ssn_grp_from_area_infa	Dictionary	Contains a list of group numbers and associated area numbers that are valid in United States Social Security numbers.
usa_ssn_indx_from_grp_infa	Dictionary	Gives Social Security Number index from the group code for the United States.
usa_state_and_locality	Dictionary	Contains a list of United States states and their constituent localities.
usa_state_and_postcode	Dictionary	Contains a list of United States states and associated postcodes.
usa_states	Dictionary	Contains a list of states in the United States.
usa_surname_ind_infa	Dictionary	Contains a list of the United States surnames.
usa_surnames_infa	Dictionary	Contains a list of surnames in the United States.
usa_tel_area_codes_infa	Dictionary	Contains a list of the United States telephone area codes.

Name	Type	Description
usa_valid_name_frmts_infa	Dictionary	Contains a list of valid person name formats in the United States.
user_defined_dates_infa	Dictionary	Contains a list of standard date formats.
v_Global_AddressValidation_Hybrid	Verifier	Verifies postal addresses in hybrid format. Hybrid fields combine street information in a single field and use discrete fields for locality, province, and postcode information.

CHAPTER 5

Data Quality bundle for Americas

The Data Quality Americas bundle is a set of pre-configured data quality assets that you can use to parse, standardize, validate, and deduplicate data from organizations in Brazil, Canada and the United States.

The assets provide solutions for general data, address data, contact data, corporate data, and product data.

Cloud Data Quality Americas bundle

You can use the rules in an out-of-the-box manner. Add one or more rules to a mapping to perform deduplication, address verification, and standardization operations on your business data.
The following table describes the Americas bundle assets:

Name	Type	Description
c_BRA_Correct_Misspellings_Name_Status	Cleanse	Standardizes misspelled names in Brazil data and adds a status value as a prefix to each standardized name.
c_BRA_Correct_Misspellings_Names	Cleanse	Corrects misspelled names in Brazil data.
c_BRA_Detect_Company_Names	Cleanse	Detects organization names in Brazil data.
c_bra_get_name_standard	Cleanse	Standardizes names in Brazil data.
c_BRA_Lbl_Lookup_Area_Code	Cleanse	Returns corresponding area names for an area code in Brazil data.
c_BRA_Lookup_Gender	Cleanse	Identifies gender of a given name in Brazil data.
c_BRA_Standardize_Company_Suffix	Cleanse	Standardizes organization names in Brazil data.
c_CAN_Get_Name_Standard	Cleanse	Replaces nicknames with standard versions of the names in Canada data.
c_dq_State_Validate	Cleanse	Validates that a given input is a valid United States state.
c_dq_Validate_County	Cleanse	Validates that a given input is a valid United States county.
c_Label_Characters	Cleanse	Assigns label "X" to a character.

Name	Type	Description
c_predirectional_std	Cleanse	Standardizes postal address terms to their equivalent short forms or abbreviations.
c_Remove_Bad_Characters	Cleanse	Removes a range of non-alphanumeric characters, such as punctuation symbols.
c_SR_FullName	Cleanse	Standardizes full names.
c_Social Security Number_Group_From_Area	Cleanse	Returns group number component in Social Security Number that corresponds to a given ZIP Code.
c_Social Security Number_Max_Index_For_Group	Cleanse	Returns index for a group.
c_US_Company_Name_Std	Cleanse	Standardizes organization names in United States data.
c_US_FullName	Cleanse	Returns label for a full name in United States data.
c_US_Get_Company_Acronym	Cleanse	Standardizes organization acronyms in United States data.
c_US_Get_Name_Standard	Cleanse	Standardizes given names in United States Data.
c_US_Lbl_OLN1	Cleanse	Returns label for last name in United States data.
c_US_Lbl_OLN2	Cleanse	Returns label for last name in United States data.
c_US_MultiPort_ContactData	Cleanse	Converts input characters in United States data to upper case.
c_US_postdirectional_std	Cleanse	Standardizes street identifiers in United States address data.
c_US_SIC_Code_Description	Cleanse	Standardizes Standard Industrial Classification (SIC) codes in United States data.
c_USA_PostCodes	Cleanse	Validates that the input data contains a valid ZIP Code in United States data.
dq_Index_for_Group	Cleanse	Returns index for group.
c_Remove_Extra_Spaces	Cleanse	Replaces multiple consecutive character spaces with a single space, and trims any leading and trailing spaces.
c_Remove_Punctuation	Cleanse	Removes all punctuations from the input data.
c_Remove_Space	Cleanse	Removes all character spaces from the input data.
c_replace_pipe_w_space	Cleanse	Replaces a pipe character with a single character space.
c_Std_Name	Cleanse	Standardizes name values by removing extraneous character spaces and punctuation symbols.

Name	Type	Description
c_Titlecase	Cleanse	Converts input characters to title case.
c_Uppercase	Cleanse	Converts input characters to upper case.
c_US_validate_area_code_label	Cleanse	Validates and assigns label to telephone area codes in Canada data.
c_Validation_Domain	Cleanse	Validates domain of an email.
c_multi_uppercase	Cleanse	Converts character case to uppercase for up to six input fields.
dedupe_BRA_CompanyName	Deduplicate	Uses organization name data to find duplicate records in Brazil data.
dedupe_BRA_Fam_Address	Deduplicate	Uses household name and address data to find duplicate records in Brazil data.
dedupe_BRA_indivdual_name_addr_match	Deduplicate	Uses person name and address data to find duplicate records in Brazil data.
dedupe_BRA_Individual_Name_and_Date	Deduplicate	Uses person name and date data to find duplicate records in Brazil data.
dedupe_BRA_Individual_Name_and_Email	Deduplicate	Uses person name and email address data to find duplicate records in Brazil data.
dedupe_BRA_Individual_Name_and_Phone	Deduplicate	Uses person name and telephone number data to find duplicate records in Brazil data.
dedupe_BRA_Match_Company_and_Address	Deduplicate	Uses organization name and address data to find duplicate records in Brazil data.
dedupe_BRA_Match_Indiv_Name_and_CPF	Deduplicate	Matches individual names and CPF for Brazil data.
dedupe_BRA_Match_Name_Personal_Data	Deduplicate	Matches personal name and data for Brazil data.
dedupe_CompanyName_Match	Deduplicate	Matches the organization name.
dedupe_US_IMO_Fam_Name_Address_Match	Deduplicate	Matches the family name and address by using Address_Part1 as key.
dedupe_US_IMO_Indiv_Name_Address_Match	Deduplicate	Matches the individual name and address by using Address Part1 as key.
dedupe_US_IMO_Match_Company_and_Address	Deduplicate	Matches the organization name and address of the United States.
dedupe_US_IMO_Personal_Name_Data_Match	Deduplicate	Matches the personal name and data.
dedupe_USA_Address_Match	Deduplicate	Matches the addresses for United States with Address Part1 as key.
bra_business_words_infa	Dictionary	Contains information about Brazilian business words.

Name	Type	Description
bra_city_from_st_area_infa	Dictionary	Contains area-wise information of Brazilian cities and states.
bra_company_suffix_std_infa	Dictionary	Contains a list of Brazilian organization name suffixes.
bra_gender_infa	Dictionary	Contains a list of Brazilian names and their genders.
bra_name_designators_infa	Dictionary	Contains a list of Brazilian name designators.
bra_name_prefix_infa	Dictionary	Contains a list of Brazilian prename prefixes.
bra_name_suffix_infa	Dictionary	Contains a list of Brazilian name suffixes.
bra_names_mispelled_infa	Dictionary	Contains a list of misspelled Brazilian names.
bra_names_mispelled_stat_infa	Dictionary	Contains a list of status values for misspelled Brazilian names.
bra_nicknames_infa	Dictionary	Contains a list of Brazilian nicknames.
can_gender_infa	Dictionary	Contains a list of genders on the basis of Canadian names.
can_nicknames_infa	Dictionary	Contains a list of nicknames in Canada.
can_postcodes	Dictionary	Contains a list of Canadian postcodes.
can_prename_gender_infa	Dictionary	Contains a list of genders on the basis of Canadian prenames.
can_provinces_infa	Dictionary	Contains list of Canadian provinces.
can_tel_area_codes_infa	Dictionary	Contains a list of Canadian telephone area codes.
usa_adrs_prfx_std_s_infa	Dictionary	Contains a list of address name standard prefixes for the United States.
usa_adrs_sufx_std_s_infa	Dictionary	Contains a list of address name standard suffixes for the United States.
usa_company_acronyms_infa	Dictionary	Contains a list of organization names and their acronyms for United States.
usa_company_names_std_infa	Dictionary	Contains a list of company name standards for the United States.
usa_company_sufx_abrv_infa	Dictionary	Contains a list of abbreviations used for company suffixes.
usa_counties_infa	Dictionary	Contains a list of counties and their possible alternate names for the United States.

Name	Type	Description
usa_gender_infa	Dictionary	Contains a list of gender based on names for the United States.
usa_name_frequencies_by_space	Dictionary	Contains a list of name frequencies by space for the United States.
usa_name_frequencies_infa	Dictionary	Contains a list of names with their frequencies for the United States.
usa_nicknames_infa	Dictionary	Contains a list of nicknames for the United States.
usa_prename_gender_infa	Dictionary	Associates gender on the basis of prename for the United States.
usa_postcodes	Dictionary	Captures the postcodes for the United States.
usa_sic_codes_infa	Dictionary	Contains the Standard Industrial Classification (SIC) codes for the United States.
usa_Social Security Number_grp_from_area_infa	Dictionary	Contains Social Security Number group from an area for the United States.
usa_Social Security Number_indx_from_grp_infa	Dictionary	Gives Social Security Number index from group code for the United States.
usa_states	Dictionary	Contains a list of states in the United States.
usa_surnames_infa	Dictionary	Contains a list of surnames in the United States.
usa_tel_area_codes_infa	Dictionary	Contains a list of telephone area codes for the United States.
domain_names_infa	Dictionary	Contains a list of domain names.
dq_av_geocodingstatus_infa	Dictionary	Contains a list of address validation status with geocoding.
dq_av_mailabilityscores_infa	Dictionary	Contains a list of 'AddressDoctor' status values, that describe the likelihood of mail delivery and a set of corresponding mailability scores in a range from 0 through 5.
dq_av_match_code_descriptions_infa	Dictionary	Contains a list of 'Match Code' output values, and identifies the process modes that return the values.
mplt_BRA_Address_Validation_Discrete	Mapplet	Validates the deliverability of Brazil addresses. Use the mapplet when you can connect the input address fields to fields on the Discrete model in the verifier asset.
mplt_BRA_Address_Validation_Discrete_w_Geo	Mapplet	Validates the deliverability of Brazil addresses and adds latitude and longitude coordinates to each valid address. Use the mapplet when you can connect the input address fields to fields on the Discrete model in the verifier asset.

Name	Type	Description
mplt_BRA_Address_Validation_Hybrid	Mapplet	Validates the deliverability of Brazil addresses. Use the mapplet when you can connect the input address fields to fields on the Hybrid model in the verifier asset.
mplt_BRA_Address_Validation_Hybrid_Geo	Mapplet	Validates the deliverability of Brazil addresses and adds latitude and longitude coordinates to each valid address. Use the mapplet when you can connect the input address fields to fields on the Hybrid model in the verifier asset.
mplt_BRA_Address_Validation_Multiline	Mapplet	Validates the deliverability of Brazil addresses. Use the mapplet when you can connect the input address fields to fields on the Multiline model in the verifier asset.
mplt_BRA_Address_Validation_Multiline_Geo	Mapplet	Validates the deliverability of Brazil addresses and adds latitude and longitude coordinates to each valid address. Use the mapplet when you can connect the input address fields to fields on the Multiline model in the verifier asset.
mplt_BRA_Company_Name_and_Address_Match	Mapplet	Finds duplicate rows of Brazilian company names and address data. The first three characters of the Soundex of the company name, and first three digits of the postcode group the address data.
mplt_BRA_Company_Name_MatchScore	Mapplet	Finds duplicate rows of Brazilian company name data. The first three letters of the Soundex of the company name group the data. Threshold frequency is 80%.
mplt_BRA_Company_Suffix_Standardization	Mapplet	Standardizes Brazilian company suffixes.
mplt_BRA_Contact_Data	Mapplet	Parses name, provides a gender, a prename (Mr. or Ms. if missing), and formal or casual greetings. Standardizes the Brazilian company name. Corrects, validates, and standardizes address, phone number, email, and personal ID.
mplt_BRA_Familyname_and_Address_Match	Mapplet	Finds duplicate rows of personal surname and Brazilian address data. The first two digits of postcode and surname group the data.
mplt_BRA_Gender_Assignment	Mapplet	Assigns gender based on the first name. Assigns last name if it has Junior or other similar words.
mplt_BRA_Given_Name_Standard	Mapplet	Returns standard name for common Brazilian nicknames. For example, it returns Robert for the input Bob.
mplt_BRA_IMO_Personal_Name_and_Data_Match	Mapplet	Finds duplicate individual names and personal data. The personal data field may have phone number, email, CPF, and other attributes with similar type. The personal data type is the group key. The personal data field sorts and groups all rows.

Name	Type	Description
mplt_BRA_Individual_Name_and_Address_Match	Mapplet	Finds duplicate rows of personal names and address data for Brazil. The NYSIIS value of surname and first three digits of the postcode group the data.
mplt_BRA_Indvidual_Name_and_Date_Match	Mapplet	Finds duplicate rows of personal names and date data for Brazil. The threshold frequency is 80%.
mplt_BRA_Individual_Name_and_Email_Match	Mapplet	Finds duplicate rows of personal names and email data for Brazil. The NYSIIS value of surname and first three digits of postcode group the data.
mplt_BRA_Individual_Name_and_CPF_Match	Mapplet	Finds duplicate rows of personal name and CPF data for Brazil. The CPF groups the data.
mplt_BRA_Individual_Name_and_Phone_Match	Mapplet	Finds duplicate rows of personal name and phone number data for Brazil. The phone number groups the data. The threshold frequency is 80%.
mplt_BRA_Personal_CPF_Validation	Mapplet	Validates CPF number with its check digits for Brazil.
mplt_BRA_Personal_CPNJ_Validation	Mapplet	Validates CPNJ number with its check digits for Brazil.
mplt_BRA_Personal_Name_Parse_Validate	Mapplet	Parses Brazilian names input in a single field into separate components. Provides information if the spelling is correct or if the data is a potential company name.
mplt_BRA_Personal_PIS_PASEP_Validation	Mapplet	Validates the number that manages government social benefits.
mplt_BRA_Personal_Voter_Registration_Validation	Mapplet	Validates digits of the Brazilian voter registration number.
mplt_BRA_Phone_Number_Parse	Mapplet	Parses phone number that follows the standard format for Brazil from a string of texts into its own field.
mplt_BRA_Phone_Number_Standardization	Mapplet	Standardizes format of the phone string to dashes without spaces for Brazil.
mplt_BRA_Phone_Validation	Mapplet	Validates area-code and length of a phone number for Brazil.
mplt_BRA_Prename_Assignment	Mapplet	Assigns prename based on gender. Uses variable to change female prename assignment value. For example, Sra or Sta.
mplt_BRA_Salutation_Assignment	Mapplet	Generates formal and casual greeting. For example, Sr John Smith gets a formal greeting of Prezado Sr Smith and the casual greeting of Prezado John. The prefix (Prezado) and suffix (,) can be changed in dq_Generate_Salutation transform.

Name	Type	Description
mplt_CAN_Address_Certification_Hybrid	Mapplet	Validates Canadian address data with input fields that map to the Hybrid model. The outputs include USA CASS Certification ports.
mplt_CAN_Address_Validation_Discrete	Mapplet	Validates the deliverability of Canada addresses. Use the mapplet when you can connect the input address fields to fields on the Discrete model in the verifier asset.
mplt_CAN_Gender_Assignment	Mapplet	Assigns a gender. For example, John Smith has M for male.
mplt_CAN_Given_Name_Standard	Mapplet	Returns standard name for common nicknames. For example, it returns Robert with Bob's input.
mplt_CAN_Personal_Name_Parsing_FML	Mapplet	Parses input names in First-Middle-Last format in discrete components.
mplt_CAN_Multi_Person_Name_Parse	Mapplet	Parses input name into proper name components for Canada.
mplt_CAN_Phone_Number_Parse	Mapplet	Parses phone number in the Canadian standard format from a string of texts into its own field.
mplt_CAN_Phone_Number_Standardization	Mapplet	Standardizes format of the phone string to have Phone_w_Extenstion and dashes without spaces.
mplt_CAN_Phone_Number_Validation	Mapplet	Validates area code and length of the Canadian phone numbers. Quality scores are: - 0 - NULL or blank input - 10 - Invalid length - 40 - Valid length, Invalid Area Code - 95 - Valid length and Area Code
mplt_CAN_Postcode_Validation	Mapplet	Validates if input string is a valid Canadian postcode.
mplt_CAN_Province_Validation	Mapplet	Validates if input string is a valid Canadian province name.
mplt_CAN_SIN_Parse	Mapplet	Parses Canadian Social Insurance Number (SIN) from a string of texts.
mplt_CAN_SIN_Standardization	Mapplet	Formats Canadian Social Insurance Number (SIN) with variable value in the expression. The Valid formats are No_Punctuation, Space or Dash. The default setting is No_Punctuation.
mplt_CAN_SIN_Validation	Mapplet	Validates Canadian Social Insurance Number (SIN).
mplt_Company_Name_Match	Mapplet	Finds duplicate rows of company name data. The first three letters of the Soundex of the company name group company name data.

Name	Type	Description
mplt_Prename_Assignment	Mapplet	Assigns a prename based on the gender. Uses a variable to change the female prename assignment value. For example, Ms or Mrs.
mplt_Salutation_Assignment	Mapplet	Generates a formal and casual greeting. For example, Mr John Smith gets a formal greeting of Dear Mr Smith, and a casual greeting of Dear John . You can change the prefix and the suffix in dq_Generate_Salutation transform.
mplt_USA_Address_Match	Mapplet	Finds duplicate rows of the United States Address Data. The first three digits of the post code group the Address Data.
mplt_USA_Address_Validation_Discrete	Mapplet	Validates the deliverability of United States addresses. Use the mapplet when you can connect the input address fields to fields on the Discrete model in the verifier asset.
mplt_USA_Address_Validation_Discrete_w_Geocoding	Mapplet	Validates the deliverability of United States addresses and adds latitude and longitude coordinates to each valid address. Use the mapplet when you can connect the input address fields to fields on the Discrete model in the verifier asset.
mplt_USA_Address_Validation_Hybrid	Mapplet	Validates the deliverability of United States addresses. Use the mapplet when you can connect the input address fields to fields on the Hybrid model in the verifier asset.
mplt_USA_Address_Validation_Hybrid_w_Geocoding	Mapplet	Validates the deliverability of United States addresses and adds latitude and longitude coordinates to each valid address. Use the mapplet when you can connect the input address fields to fields on the Hybrid model in the verifier asset.
mplt_USA_Address_Validation_Multiline	Mapplet	Validates the deliverability of United States addresses. Use the mapplet when you can connect the input address fields to fields on the Multiline model in the verifier asset.
mplt_USA_Address_Validation_Multiline_w_Geocoding	Mapplet	Validates the deliverability of United States addresses and adds latitude and longitude coordinates to each valid address. Use the mapplet when you can connect the input address fields to fields on the Multiline model in the verifier asset.
mplt_USA_Company_Name_Standardization	Mapplet	Standardizes company name and provides its acronym if possible.
mplt_USA_Contact_Data	Mapplet	Parses name and provides missing gender and prename (Mr. or Ms.). Provides formal or casual greetings. Standardizes company name. Corrects, validates, and standardizes the United States addresses. Standardizes and validates phone, email, and Social Security Number.

Name	Type	Description
mplt_USA_County_Validation	Mapplet	Checks if the input string is a valid county name in any state.
mplt_USA_Gender_Assignment	Mapplet	Assigns gender. For example, a name of John Smith has M for male.
mplt_USA_Given_Name_Standard	Mapplet	Returns standard name for common nicknames. For example, the input Bob returns Robert.
mplt_USA_IMO_Company_Name_and_Address_Match	Mapplet	Finds duplicate rows of company name and address data for the United States.
mplt_USA_IMO_Familyname_and_Address_Match	Mapplet	Finds duplicate rows of personal surname and address data for the United States. The first three digits of the postcode and the surname group the rows.
mplt_USA_IMO_Individual_Name_and_Address_Match	Mapplet	Finds duplicate rows of personal name and address data for the United States with Identity Match option.
mplt_USA_IMO_Personal_Name_and_Data_Match	Mapplet	Finds duplicate individual names and personal data for the United States. The personal data field may have any type of data such as phone number, email, and Social Security Number, but should be of the same type. The personal data type is the group key that sorts and groups all rows.
mplt_USA_Multi_Person_Name_Parse	Mapplet	Parses the United States input name into proper name components.
mplt_USA_Personal_Name_Parse_Validation	Mapplet	Uses gender-specific score to determine the status of a known gender for any name. Accepts M and F as inputs for male and female genders respectively. Uses the highest male or female scores for unknown gender to determine the status. Calculates the probable gender based on first name input and provides a confidence score based on the frequency of male or female name occurrences. The genders have a score if the probability of the male or female gender is 70% or more. Unknown genders have a confidence score of zero.
mplt_USA_Personal_Name_Parse_and_Standardize_LFM_Prebuilt	Mapplet	Uses prebuilt parser that parses name in Last-First-Middle format into discrete components. Assigns prenames and standardizes the prename, postname, and title cases for all the components. Builds a full name from discrete components.
mplt_USA_Personal_Name_Parsing_FML	Mapplet	Parses United States input names in First-Middle-Last format in discrete components.
mplt_USA_Personal_Name_Parsing_LFM	Mapplet	Parses a United States name in Last-First-Middle format in discrete components.
mplt_USA_Phone_Number_Parse	Mapplet	Parses phone number in the United States format from a string of texts in its own field.

Name	Type	Description
mplt_USA_Phone_Number_Standardization	Mapplet	Standardizes format of United States phone string to standard dashes without spaces.
mplt_USA_Phone_Number_Validation	Mapplet	Validates area code for the United States phone numbers. The quality scores are: - 0 - NULL or blank input - 10 - Invalid length - 40 - Valid length, Invalid Area Code - 95 - Valid length and Area Code
mplt_USA_Phone_Parse_Standardize_Validate	Mapplet	Parses phone number for the United States from a string of texts. Validates that area code is valid for the United States, and returns phone number in three different standard formats for valid codes. Returns the status value (Valid or Invalid) and the status description.
mplt_USA_Phone_w_Extension_Parse	Mapplet	Parses phone number that can have an extension as per the United States standard format from a string of texts into its own field.
mplt_USA_SIC_Code_Validation	Mapplet	Validates Standard Industrial Classification (SIC) codes for the United States. Returns the status of valid and invalid codes, the descriptions, and the short explanations.
mplt_USA_Social Security Number_Parse	Mapplet	Parses Social Security Number patterns for the United States with or without dashes from a larger string of texts.
mplt_USA_Social Security Number_Parse_Standardize_and_Validate	Mapplet	Parses Social Security Number patterns for the United States with or without dashes from a larger string of texts. Formats and validates Social Security Number. Format options include without punctuation, with punctuation, and spaces. Change the variable port Social Security Number_Format in dq_Social Security Number_Format transform as desired. Default output is without punctuation.
mplt_USA_Social Security Number_Standardization	Mapplet	Formats input Social Security Number for the United States. Formats include without punctuation, with punctuation, and with spaces. Change the variable port Social Security Number_Format in the dq_Social Security Number_Format transform as desired. The default output is without punctuation.
mplt_USA_Social Security Number_Validation	Mapplet	Checks if area (first 3 numbers) and group (middle 2 numbers) are valid combinations for all the 0s. Does not check if Social Security Number is an issued number. Accurately validates if Social Security Numbers exist before June 25, 2011 due to Social Security Number Randomization.

Name	Type	Description
mplt_USA_Social Security Number_Validation_post_June2011	Mapplet	Validates Social Security Number for length, numeric values and known minimum or maximum values of the area, group, and serial number sections. Does not check if Social Security Number is an issued number. Does not validate group and area number combination because Social Security Numbers exist before June 25, 2011 due to Social Security Number Randomization.
mplt_USA_State_Validation	Mapplet	Validates input string as a state name for the United States.
mplt_USA_ZIPCode_Validation	Mapplet	Validates input string as a 5-digit United States zipcode.
mplt_Email_Validation	Mapplet	Validates email format. Does not check if the email addresses are active. Returns valid or invalid emails.
mplt_Remove_Extra_Spaces	Mapplet	Replaces multiple consecutive spaces with a single character space and trims leading and trailing spaces.
mplt_Remove_Non_Numbers	Mapplet	Removes all characters that are not digits or numbers from the input strings.
mplt_Remove_Punctuation	Mapplet	Removes all punctuation occurrences from the input strings.
mplt_Remove_Punctuation_and_Space	Mapplet	Removes all punctuations and spaces from the input strings, and returns the digits and letters.
mplt_Replace_Limited_Punct_with_Space	Mapplet	Replaces certain punctuations with a single character space. The punctuations include: /, \, ., !, _ and two, three, and four consecutive spaces.
mplt_TitleCase	Mapplet	Converts case of the input text to title case. The first character of each word is in upper case, while the remaining characters are in lower cases.
mplt_UpperCase	Mapplet	Changes cases to upper for all input data.
p_Assign_Firstname_Gender_Can	Parse	Assigns gender to first name for Canada data.
p_Assign_Midname_Gender_Can	Parse	Assigns gender to middle name for Canada data.
p_Assign_Prename_Gender_can	Parse	Assigns gender to prename for Canada data.
p_BRA_parse_NameDesignator	Parse	Parses name designator for Brazil data.
p_BRA_Parse_PostName	Parse	Parses postname for Brazil data.
p_BRA_Parse_Prename	Parse	Parses prename for Brazil data.
p_can_postcode_validation	Parse	Validates postcodes for Canada data.

Name	Type	Description
p_can_Province_Validation	Parse	Parses Canadian provinces and validates the inputs.
p_dq_Label_Noise	Parse	Removes characters in a phone number.
p_parse_bra_phone_number	Parse	Parses phone number for Brazil data from a string input.
p_parse_can_phone_number	Parse	Parses phone number for Canada data from a string input.
p_parse_CAN_SIN	Parse	Parses Social Insurance Number for Canada.
P_Parse_Name_Components_FML_Can	Parse	Parses name components for Canada.
p_Parse_Name_Components_FML_USA	Parse	Parses inputs into First Name, Middle Name, and Last Name for the United States.
P_Parse_Name_Components_LFM_USA	Parse	Parses inputs into Last Name, Middle Name, and First Name for the United States.
p_Parse_Social Security Number	Parse	Parses Social Security Number from a string.
p_USA_Assign_metric_to_matched_values	Parse	Assigns metrics for the United States based on reference table.
p_usa_gender_assignemnt_FN	Parse	Assigns gender based on first name.
p_usa_gender_assingment_MN	Parse	Assigns gender based on middle name.
p_USA_gender_Assingment_Prename	Parse	Assigns gender based on prename.
p_USA_Phone	Parse	Parses phone number for the United States from a string input.
p_usa_phone_w_extension	Parse	Parses phone number with extension for the United States from a larger string.
p_USA_Social Security Number	Parse	Parses Social Security Number for the United States in XXX-XX-XXXX format.
p_AV_MailabilityScores_infa	Parse	Parses malleability score.
p_FN_metrics_to_ind_fields	Parse	Parses metrics.
p_Parse_Assign_GeocodingStatus_Desc	Parse	Parses geocoding status description.
p_Parse_CAN_FullName	Parse	Parses full name in multi-name format for Canada data.
p_Parse_CAN_validate_area_code	Parse	Validates area code for Canada data.
p_Parse_US_FullName	Parse	Parses full name in multi-name format for the United States data.

Name	Type	Description
p_SN_metrics_to_ind_fields	Parse	Parses surname metrics from a single input into discrete fields.
rs_can_postcode_validation	Rule Specification	Validates postcodes for Canada data.
rs_can_province_validation	Rule Specification	Validates province names for Canada data.
rs_USA_SIC_Validation	Rule Specification	Validates Standard Industrial Classification (SIC) codes in the United States data.
rs_Title_Case	Rule Specification	Converts a given input to title case.
rs_validate_area_code	Rule Specification	Identifies null entries in a column of telephone area code data.
v_BRA_Address_Val_Hybrid_Address	Verifier	Verifies Brazilian addresses with input fields that map to the Hybrid model.
v_BRA_AddressValidation_Hyb_w_Geo	Verifier	Verifies Brazilian addresses with input fields that map to the Hybrid model, and adds geocodes to the output.
v_BRA_AddressValidation_Multiline	Verifier	Verifies Brazilian addresses with input fields that map to the Multiline model.
v_BRA_Parse_Multiline_Address_w_Geo	Verifier	Verifies Brazilian addresses with input fields that map to the Multiline model, and adds geocodes to the output.
v_BRA_Validation_Discrete_Address	Verifier	Verifies Brazilian addresses with input fields that map to the Discrete model.
v_BRA_Validation_Discrete_Address_Geo	Verifier	Verifies Brazilian addresses with input fields that map to the Discrete model, and adds geocodes to the output.
v_CAN_AddressVerification_Hybrid	Verifier	Validates Canadian addresses with input fields that map to the Hybrid model.
v_CAN_AddressValidation_Discrete	Verifier	Verifies Canadian addresses with input fields that map to the Discrete model.
v_USA_AddressValidation_Hyb_w_Geo	Verifier	Verifies United States address with input fields that map to the Hybrid model, and adds geocodes to the output.
v_USA_AddressValidation_Hybrid	Verifier	Verifies United States addresses with input fields that map to the Hybrid model.

Name	Type	Description
v_USA_AddressValidation_Geo_Multiline	Verifier	Verifies United States addresses with input fields that map to the Multiline model, and adds geocodes to the output.
v_USA_AddressValidation_Multiline	Verifier	Verifies United States addresses with input fields that map to the Multiline model.
v_USA_Validation_Discrete_Address	Verifier	Verifies United States addresses with input fields that map to the Discrete model.
v_USA_Validation_Discrete_Address_w_Geo	Verifier	Verifies United States addresses with input fields that map to the Discrete model, and adds geocodes to the output.

CHAPTER 6

Data Quality bundle for United Kingdom

Use the Data Quality bundle for United Kingdom to accelerate the configuration and deployment of data quality solutions in United Kingdom organizations. The bundle includes mapplets and other assets that define data standardization, deduplication, address verification, and parsing operations for United Kingdom data.

The following table lists the assets contained in the Data Quality bundle for United Kingdom:

Name	Asset Type	Description
c_GBR_Get_Company_Acronym	Cleanse	Standardizes acronyms for United Kingdom company names.
c_GBR_Get_Company_Name_Std	Cleanse	Standardizes United Kingdom company names.
c_GBR_given_name_standard	Cleanse	Standardizes given names and removes extraneous character spaces.
c_GBR_Remove_Intl_Dial_Code	Cleanse	Removes variations of the United Kingdom international dialing code from the start of the input string. For example, `+44` and `0044`
c_GBR_Standardize_Company_Suffix	Cleanse	Standardizes United Kingdom company name suffixes.
c_Remove_Extra_Spaces	Cleanse	Replaces multiple consecutive spaces with a single space and trims leading and trailing spaces.
c_Remove_Intl_call_Prefix	Cleanse	Removes the international dialing prefix from telephone numbers.
c_remove_labels	Cleanse	Removes the characters `X` and `9`.
c_Remove_Leading_Zero	Cleanse	Removes the first character from a number if the character is a zero.
c_Remove_Numbers	Cleanse	Removes numbers from a string.

Name	Asset Type	Description
c_Remove_Period_Parentheses	Cleanse	Removes all periods and left and right parentheses.
c_Remove_Punctuation	Cleanse	Removes all punctuation from the input field.
c_Remove_Space	Cleanse	Removes all occurrences of a character space from the input field.
c_replace_ltd_punct_w_space	Cleanse	Replaces a limited set of punctuation symbols with character spaces.
c_replace_numbers_alpha_w_9X	Cleanse	Replaces all digits with `9` and all alphabetical characters with `X`.
c_SR_FullName	Cleanse	Standardizes full names.
c_Titlecase	Cleanse	Converts the input text to title case.
c_Uppercase	Cleanse	Converts the input text to uppercase.
c_Validation_Domain	Cleanse	Validates the domain of an email address.
dedupe_GBR_CompanyName_Postcode_Match	Deduplicate	Identifies duplicate records in United Kingdom data based on company names and postal codes.
dedupe_GBR_Familyname_NINO_Match	Deduplicate	Identifies duplicate records in United Kingdom data based on surnames and National Insurance Numbers (NINO).
dedupe_GBR_Familyname_Postcode_Match	Deduplicate	Identifies duplicate records in United Kingdom data based on surnames and United Kingdom postal codes.
dedupe_GBR_FN_3CharSN_DOB_Zip_Match	Deduplicate	Identifies duplicate records in United Kingdom data based on the following data: - First name - The first three characters in the surname - Date of birth - Postal code

Name	Asset Type	Description
dedupe_GBR_FN_SN_2EDOB_Postcode_Match	Deduplicate	Identifies duplicate records in United Kingdom data based on the following data: - Person name - Any two date of birth elements, such as month and year - United Kingdom postal code
dedupe_GBR_FN_SN_DOB_Postcode_Match	Deduplicate	Identifies records based on the following data: - Person name - Date of birth - Postal code
dedupe_GBR_IMO_Comp_Name_Addr_Match	Deduplicate	Identifies duplicate records in United Kingdom data based on company names and addresses.
dedupe_GBR_Individual_Name_and_Email	Deduplicate	Identifies duplicate records in United Kingdom data based on person names and email address data.
Dedupe_GBR_Individual_Name_and_NINO	Deduplicate	Identifies duplicate records in United Kingdom data based on person names and National Insurance Numbers (NINO).
Dedupe_GBR_Name_Postcode_Match	Deduplicate	Identifies duplicate records in United Kingdom data based on person names and postcode data.
domain_names_infa	Dictionary	Contains a list of domain names.
dq_av_geocodingstatus_infa	Dictionary	Contains geocoding status values and their descriptions.
dq_av_match_code_descriptions_infa	Dictionary	Contains alphanumeric status values that indicate the outcome of the verification operation for an address. Each status value has a corresponding text description.
gbr_bank_sort_codes_infa	Dictionary	Contains United Kingdom bank sort codes.
gbr_company_acronyms_infa	Dictionary	Contains United Kingdom company name acronyms.
gbr_company_name_sample	Dictionary	Contains sample United Kingdom company names.
gbr_company_names_infa	Dictionary	Contains a list of United Kingdom company names.

Name	Asset Type	Description
gbr_company_sufx_abrv_infa	Dictionary	Contains company suffix abbreviations in use in the United Kingdom.
gbr_gender_infa	Dictionary	Contains a list of names and corresponding genders in the United Kingdom.
gbr_nicknames_infa	Dictionary	Contains a list of United Kingdom nicknames.
gbr_prename_gender_infa	Dictionary	Contains name prefixes and the gender associated with each prefix.
gbr_tel_area_codes_infa	Dictionary	Contains a list of United Kingdom telephone area codes.
title_case_excptn_infa	Dictionary	Contains a list of titles that precede personal names.
lbl_GBR_Area_Code_2Digits	Labeler	Labels two-digit United Kingdom telephone area codes.
lbl_GBR_Area_Code_3Digits	Labeler	Labels three-digit United Kingdom telephone area codes.
lbl_GBR_Area_Code_4Digits	Labeler	Labels four-digit United Kingdom telephone area codes.
lbl_GBR_NINO	Labeler	Labels United Kingdom National Insurance Numbers (NINO).
mplt_Assign_DQ_GeocodingStatus_Description	Mapplet	Assigns the text description of a Geocoding Accuracy Code value that a verifier asset generates.
mplt_Assign_DQ_Match_Code_Description	Mapplet	Assigns the text description of a Verification Status Code value that a verifier asset generates.
mplt_Email_Validation	Mapplet	Validates the format of email addresses. The mapplet does not verify that the email addresses are accurate or active. Returns Valid or Invalid.
mplt_GBR_Address_Validation_Discrete	Mapplet	Validates the deliverability of United Kingdom addresses. Use the mapplet when you can connect the input address fields to fields on the Discrete model in the verifier asset.

Name	Asset Type	Description
mplt_GBR_Address_Validation_Discrete_w_Geocoding	Mapplet	Validates the deliverability of United Kingdom addresses and adds latitude and longitude coordinates to each valid address. Use the mapplet when you can connect the input address fields to fields on the Discrete model in the verifier asset.
mplt_GBR_Address_Validation_Hybrid	Mapplet	Validates the deliverability of United Kingdom addresses. Use the mapplet when you can connect the input address fields to fields on the Hybrid model in the verifier asset.
mplt_GBR_Address_Validation_Hybrid_Geocoding	Mapplet	Validates the deliverability of United Kingdom addresses and adds latitude and longitude coordinates to each valid address. Use the mapplet when you can connect the input address fields to fields on the Hybrid model in the verifier asset.
mplt_GBR_Address_Validation_Multiline	Mapplet	Validates the deliverability of United Kingdom addresses. Use the mapplet when you can connect the input address fields to fields on the Multiline model in the verifier asset.
mplt_GBR_Address_Validation_Multiline_w_Geocoding	Mapplet	Validates the deliverability of United Kingdom addresses and adds latitude and longitude coordinates to each valid address. Use the mapplet when you can connect the input address fields to fields on the Multiline model in the verifier asset.
mplt_GBR_Bank_Account_Parse	Mapplet	Parses an eight-digit string as a United Kingdom bank account number from a string.
mplt_GBR_Bank_Account_Validation	Mapplet	Validates United Kingdom bank account numbers. The mapplet checks if the input is numeric and if the value is eight characters long.
mplt_GBR_Bank_Sort_Code_Parse	Mapplet	Parses six-digit numeric strings as United Kingdom bank sort codes. The mapplet parses strings of numbers in the following formats: - Consecutive numbers (999999) - Numbers delimited with a dash (99-99-99)
mplt_GBR_Bank_Sort_Code_Standardize	Mapplet	Standardizes a United Kingdom bank sort code to the format NN-NN-NN.

Name	Asset Type	Description
mplt_GBR_Bank_Sort_Code_Validation	Mapplet	Validates the format and length of United Kingdom bank sort codes that are standardized to the dash-delimited format (99-99-99). The mapplet returns a Status field that describes the validity of the sort code and a Validation Note field that explains the status. If the sort code prefix matches a known assignment for a United Kingdom bank, the validation note includes the bank name.
mplt_GBR_Company_Name_Postcode_Match	Mapplet	Identifies duplicate records in United Kingdom data based on company names and postal codes. The mapplet generates group keys from the postal codes.
mplt_GBR_Company_Name_Standardization	Mapplet	Standardizes a company name and provides the acronym for the name if it is possible to do so.
mplt_GBR_Contact_Data	Mapplet	Parses, standardizes, and validates United Kingdom contact data, such as addresses, telephone numbers, and National Insurance Numbers (NINO).
mplt_GBR_Driver_Number_Parse	Mapplet	Parses strings that match the format of United Kingdom driver's license numbers from a string.
mplt_GBR_Driver_Number_Validation	Mapplet	Validates United Kingdom driver's license numbers based on the requirements of the United Kingdom Government Data Standards Catalogue.
mplt_GBR_Familyname_NINO_Match	Mapplet	Identifies duplicate records in United Kingdom data based on surnames and National Insurance Numbers (NINO). The mapplet generates group keys from the NINO data.
mplt_GBR_Familyname_Postcode_Match	Mapplet	Identifies duplicate records in United Kingdom data based on surnames and United Kingdom postal codes. The mapplet generates group keys from the postal code data.

Name	Asset Type	Description
mplt_GBR_Firstname_3CharsSurname_DOB_Postcode_Match	Mapplet	Identifies duplicate records in United Kingdom data based on the following data: - First name - The first three characters in the surname - Date of birth - Postal code The mapplet generates group keys from the postal code data.
mplt_GBR_Firstname_Surname_2ElementsDOB_Postcode_Match	Mapplet	Identifies duplicate records in United Kingdom data based on the following data:- - Person names - Any two date of birth elements, such as month and year - United Kingdom postal code The mapplet generates group keys from the postal code data.
mplt_GBR_Firstname_Surname_DOB_Postcode_Match	Mapplet	Identifies duplicate records based on the following data: - Person names - Date of birth - Postal code The mapplet generates group keys from the postal code data.
mplt_GBR_Gender_Assignment	Mapplet	Assigns gender to first names. The mapplet returns M for male names, F for female names, and U if the gender is unknown. For example, the mapplet assigns the name John Smith a gender of M for male.
mplt_GBR_Given_Name_Standard	Mapplet	Generate given names from United Kingdom nicknames. For example, the mapplet standardizes the nickname Bob to the given name Robert.
mplt_GBR_IMO_Company_Name_Address_Match	Mapplet	Identifies duplicate records in United Kingdom data based on company names and addresses.
mplt_GBR_Individual_Name_and_Email_Match	Mapplet	Identifies duplicate records in United Kingdom data based on person names and email address data. The mapplet generates group keys from the email address data.

Name	Asset Type	Description
mplt_GBR_Individual_Name_and_NINO_Match	Mapplet	Identifies duplicate records in United Kingdom data based on person names and National Insurance Numbers (NINO). The mapplet generates group keys from the NINO data.
mplt_GBR_Individual_Name_and_Postcode_Match	Mapplet	Identifies duplicate records in United Kingdom data based on person names and postal code data. The mapplet generates group keys from the postal code data.
mplt_GBR_Multi_Person_Name_Parse	Mapplet	Parses person name values into separate fields. The mapplet creates fields for values such as title, first name, middle name, and surname.
mplt_GBR_NHS_Number_Parse	Mapplet	Parses National Health Service (NHS) numbers from a string.
mplt_GBR_NHS_Number_Standardize	Mapplet	Standardizes National Health Service (NHS) numbers into a standard format (999 999 9999). The mapplet requires that the input is a 10-digit string. If the input value does not yield a 10-digit value, the mapple returns the input value.
mplt_GBR_NHS_Number_Validate	Mapplet	Validates National Health Service (NHS) numbers based on the check digit for each number. The mapplet requires that the input is a 10-digit string.
mplt_GBR_NINO_Conformity_Check	Mapplet	Validates that the input value conforms to a standard pattern for United Kingdom National Insurance Numbers (NINO). The mapplet does not verify that a NINO is accurate or active.
mplt_GBR_NINO_Parse	Mapplet	Parses United Kingdom National Insurance Numbers (NINO) from strings. The mapplet returns the NINO and also returns a string that contains the input text with the NINO removed.

Name	Asset Type	Description
mplt_GBR_NINO_Standardization	Mapplet	Standardizes United Kingdom National Insurance Numbers (NINO) into the two most typical formats. The mapplet returns the following formats, where C represents alphabetic characters and N represents numerals: - CC NN NN NN C - CCNNNNNNC The mapplet formats all alphabetic characters as uppercase.
mplt_GBR_NINO_Validation	Mapplet	Validates a United Kingdom National Insurance Number (NINO).
mplt_GBR_Passport_Number_MR_Parse	Mapplet	Parses United Kingdom passport numbers that appear in the machine-readable extended format.
mplt_GBR_Passport_Number_Parse	Mapplet	Parses nine-digit strings that conform to the United Kingdom passport number format that the Government Data Standards Catalogue specifies.
mplt_GBR_Passport_Number_Validation	Mapplet	Validates nine-digit strings that conform to the United Kingdom passport number format that the Government Data Standards Catalogue specifies.
mplt_GBR_Personal_Name_Parsing_FML	Mapplet	Parses person name values into separate fields. The mapplet creates the fields in the following sequence: - First name, middle name, last name
mplt_GBR_Personal_Name_Parsing_LFM	Mapplet	Parses the person name values into separate fields. The mapplet creates the fields in the following sequence: - Last name, first name, middle name

Name	Asset Type	Description
mplt_GBR_Phone_Number_Parse	Mapplet	Parses a United Kingdom telephone number from a string. The mapplet recognizes telephone numbers that use leading zeros, the +44 international dialing code, and extensions that begin with the hash symbol. The mapplet processes the following punctuation symbols: the plus sign, parentheses, and the hash symbol. Before you run the mapplet, remove all other punctuation, including double spaces.
mplt_GBR_Phone_Number_Standardization	Mapplet	Standardizes United Kingdom telephone numbers to international and local dialing formats. The mapplet recognizes telephone numbers that use leading zeros, the +44 international dialing code, and extensions that begin with the hash symbol.
mplt_GBR_Phone_Number_Validation	Mapplet	Validates the area code and length of United Kingdom telephone numbers. The mapplet can also return the region that the area code identifies.
mplt_GBR_Postcode_Parse	Mapplet	Parses United Kingdom postal codes from a string.
mplt_GBR_Postcode_Standardise	Mapplet	Standardizes United Kingdom postal codes. The mapplet requires that the input follows predefined formats. The mapplet standardizes inputs that match the following patterns: - A9 9AA - A99 9AA - AA9 9AA - AA99 9AA - A9A 9AA - AA9A 9AA - GIR 0AA If the input value does not match the above patterns, the mapplet returns the input value.
mplt_GBR_Postcode_Validate	Mapplet	Validates United Kingdom postal codes. If the mapplet does not find a postal code, it verifies whether the input follows a standard United Kingdom postal code pattern.
mplt_LowerCase	Mapplet	Converts input text to lowercase characters.

Name	Asset Type	Description
mplt_Prename_Assignment	Mapplet	Assigns a name prefix based on gender. Includes a variable that can change a female prefix. For example, Ms. or Mrs.
mplt_Remove_Extra_Spaces	Mapplet	Replaces multiple consecutive character spaces with a single space, and trims leading and trailing spaces.
mplt_Remove_Leading_Zero	Mapplet	Removes a leading zero from a number.
mplt_Remove_Period_Parantheses	Mapplet	Removes all occurrences of a period and left and right parentheses.
mplt_Remove_Punctuation	Mapplet	Removes all punctuation symbols.
mplt_Remove_Punctuation_and_Space	Mapplet	Removes all punctuation symbols and character spaces.
mplt_Remove_Space	Mapplet	Removes all occurrences of a character space.
mplt_Replace_Limited_Punct_with_Space	Mapplet	Replaces punctuation symbols, including a forward slash, backslash, exclamation point, period, or underscore character, with a character space. Also replaces instances of two, three, or four consecutive character spaces with a single space.
mplt_Salutation_Assignment	Mapplet	Generates formal and casual greetings from prefixes and names. For example, when input data contains "Mr. John Smith," the mapplet generates the formal greeting "Dear Mr. Smith," and the casual greeting "Dear John,". You can change the prefix and punctuation by editing the variables in the dq_Generate_Salutation Expression transformation.
mplt_UpperCase	Mapplet	Changes alphabetic characters to uppercase.
p_Assign_Firstname_Gender	Parse	Assigns a gender value based on the first name.
p_Assign_Midname_Gender	Parse	Assigns a gender value based on the middle name.
p_Assign_Prename_Gender	Parse	Assigns a gender value based on the name prefix.

Name	Asset Type	Description
p_GBR_driver_number	Parse	Parses a United Kingdom driver's license number from a string.
p_GBR_NINO	Parse	Parses a United Kingdom National Insurance Number (NINO) using a regular expression.
p_GBR_ParseFullName	Parse	Parses a full name from the input field.
p_GBR_Postcode_Validate_Format	Parse	Parses a United Kingdom postcode that matches a United Kingdom postcode format.
p_GBR_Postcode	Parse	Parses a United Kingdom postcode using a regular expression.
p_last_alphabet	Parse	Parses the last alphabetical character from a string.
p_NHS_Number	Parse	Parses a United Kingdom National Health Service (NHS) number using regular expressions.
p_Parse_Assign_GeocodingStatus_Desc	Parse	Parses geocoding status descriptions.
p_Parse_GBR_Acct_Number	Parse	Parses a United Kingdom account number.
p_Parse_GBR_Area_Code_2Digit	Parse	Parses two-digit telephone area codes for the United Kingdom.
p_Parse_GBR_Area_Code_3Digit	Parse	Parses three-digit telephone area codes for the United Kingdom.
p_Parse_GBR_Area_Code_4Digit	Parse	Parses four-digit telephone area codes for the United Kingdom.
p_Parse_GBR_Bank_Sort_Code	Parse	Parses United Kingdom bank sort codes.
p_Parse_GBR_Match_Bank_Sort_Codes	Parse	Parses a United Kingdom bank sort code from the input data when an input value matches a value in a dictionary.
p_Parse_GBR_Name_Components_FML	Parse	Parses name components for United Kingdom: F
p_Parse_GBR_Name_Components_LFM	Parse	Parses name components for United Kingdom: L
p_Parse_GBR_Passport_Number	Parse	Parses United Kingdom passport numbers.

Name	Asset Type	Description
p_Parse_GBR_Passport_Number_MR	Parse	Parses machine-readable United Kingdom passport numbers.
p_Parse_GBR_Phone	Parse	Parses United Kingdom telephone numbers.
p_UK_Postcode_Validate	Parse	Parses United Kingdom postcodes using a regular expression.
p_Word	Parse	Extracts the first alphabetical string from the input field where the alphabetical string is defined by a leading or trailing character space. For example, extracts `Check` from `Check Test`, `testword` from `testword 123`, or `Check` from `123 Check`.
rs_Assign_DQ_Match_Code_Description	Rule Specification	Returns the text description of a Verification Status Code value that a verifier asset generates.
av_GBR_AddressValidation_Discrete	Verifier	Verifies the deliverability of United Kingdom addresses. Use the asset when you can connect the input address fields to fields on the Discrete model.
av_GBR_AddressValidation_Discrete_w_Geo	Verifier	Verifies the deliverability of United Kingdom addresses and adds latitude and longitude coordinates to each valid address. Use the asset when you can connect the input address fields to fields on the Discrete model.
av_GBR_AddressValidation_Hybrid	Verifier	Verifies the deliverability of United Kingdom addresses. Use the asset when you can connect the input address fields to fields on the Hybrid model.
av_GBR_Hybrid_w_Gecoding	Verifier	Verifies the deliverability of United Kingdom addresses and adds latitude and longitude coordinates to each valid address. Use the asset when you can connect the input address fields to fields on the Hybrid model.

Name	Asset Type	Description
av_GBR_AddressValidation_Multiline_Geo	Verifier	Verifies the deliverability of United Kingdom addresses and adds latitude and longitude coordinates to each valid address. Use the asset when you can connect the input address fields to fields on the Multiline model.
av_GBR_Parse_Multiline_Address	Verifier	Parses unstructured United Kingdom addresses into address elements. Use the asset when you can connect the input address fields on the Multiline model.

CHAPTER 7

Data Quality bundle for Australia and New Zealand

Use the Data Quality bundle for Australia and New Zealand to accelerate the configuration and deployment of data quality solutions in Australia and New Zealand organizations. The bundle includes mapplets and other assets that define data standardization, deduplication, address verification, and parsing operations for Australia and New Zealand data.

The following table lists the assets contained in the Data Quality bundle for Australia and New Zealand:

Name	Asset Type	Description
abn_acn_exception_values	DICTIONARY	Contains a list of values that cannot be used as a valid Australian Business Number (ABN) or Australian Company Number (ACN).
aus_bsb_codes	DICTIONARY	Contains a list of Australian bank branch codes.
aus_company_name_std_infa	DICTIONARY	Contains a sample of company names and alternatives.
aus_gender_infa	DICTIONARY	Contains a list of given names with associated gender. Includes "M" for male names, "F" for female names, and "U" if the gender is unknown.
aus_nicknames_infa	DICTIONARY	Contains a list of names and alternatives.
aus_prename_gender_infa	DICTIONARY	Contains a list of courtesy titles and the gender associated with the titles.
aus_telephone_prefix	DICTIONARY	Contains a list of Australian telephone prefix numbers from Australian Government Federal Register of Legislation.
c_aus_areacode_region	CLEANSE	Checks the input data against a list of Australian telephone prefix numbers from Australian Government Federal Register of Legislation and returns `VALID` if the data values match.
c_aus_get_company_name_std	CLEANSE	Standardizes company name for Australia.
c_aus_remove_intl_dialling_prefix	CLEANSE	Removes variations of the Australian international dialing code from the start of the input string. For example: `+61, 0061, 61`

Name	Asset Type	Description
c_av6_assign_dq_status_code_description	CLEANSE	Returns the text description of a Match Code value that a verifier asset generates.
c_get_name_standard	CLEANSE	Replaces Australian nick names with the defined valid column.
c_merge_fields	CLEANSE	Merges the Overflow, Before, After and Unparsed fields from the parse asset.
c_merge_two_fields	CLEANSE	Merges two fields.
c_multi_uppercase	CLEANSE	Converts character case to uppercase for up to six input fields.
c_nzl_areacode_region	CLEANSE	Returns the name of the area code from the area code dialing prefix.
c_nzl_phone_line	CLEANSE	Removes the area codes from a New Zealand phone number.
c_nzl_remove_intl_dialling_prefix	CLEANSE	Removes the New Zealand international dialing prefixes.
c_nzl_standardize	CLEANSE	Replaces multiple zeros at the start of a string with a single zero.
c_remove_labelled_character	CLEANSE	Replaces an X, _ and S with @.
c_Remove_Non_Digit_Chars	CLEANSE	Removes @ and + symbols.
c_remove_non_numbers	CLEANSE	Removes non-numbers from a string.
c_Remove_Punctuation	CLEANSE	Removes all occurrences of punctuation from the input field.
c_Remove_Space	CLEANSE	Removes all occurrences of a character space from the input field.
c_remove_symbols_chars	CLEANSE	Removes symbols and letters.
c_remove_x	CLEANSE	Removes X or x from the start of an input.
c_Replace_Limited_Punctuation_w_Space	CLEANSE	Replaces any forward slash, backslash, exclamation point, period, or underscore character with a character space. Also replaces instances of multiple character spaces with a single space.
c_standardise_area_code	CLEANSE	Replaces the area code with the region.
c_Uppercase	CLEANSE	Converts the input text to uppercase.

Name	Asset Type	Description
dedupe_aus_familyname_and_address_match	DEDUPLICATE	Identifies duplicate records in Australian data based on similarities in family name and address data. The operation uses Address Part1 as the index key.
dedupe_aus_individual_name_and_address	DEDUPLICATE	Identifies duplicate records in Australian data based on similarities in person name and address data.
dedupe_aus_individual_name_and_date	DEDUPLICATE	Identifies duplicate records in Australian data based on similarities in person name and date data.
dedupe_aus_individual_name_and_email	DEDUPLICATE	Identifies duplicate records in Australian data based on similarities in person name and email address data. The operation uses Individual as the objective type.
dedupe_aus_individual_name_date	DEDUPLICATE	Identifies duplicate records in Australian data based on similarities in person name and date data.
dedupe_aus_individual_name_match	DEDUPLICATE	Identifies duplicate records in Australian data based on similarities in person name data.
dedupe_aus_individual_name_phone	DEDUPLICATE	Identifies duplicate records in Australian data based on person name and telephone number data.
dedupe_aus_individual_name_tfn_match	DEDUPLICATE	Identifies duplicate records in Australian data based on person name and Tax File Numbers (TFN).
dedupe_aus_personal_name_data	DEDUPLICATE	Identifies duplicate records in Australian data based on person name and personal data.
dq_av6_match_code_descriptions_infa	DICTIONARY	Contains alphanumeric status values that indicate the outcome of the verification operation for an address. Each status value has a corresponding text description.
lbl_abn_acn_exception_values	LABELER	Checks if the input value is in the list of Australian Business number exception values and marks the value as invalid if present and valid if not present.
lbl_all_characters	LABELER	Labels English alphabets, character spaces, and symbols.
lbl_aus_bank_code	LABELER	Labels Australian bank codes.
lbl_australia_bank_acct_number	LABELER	Labels Australian bank account numbers.
lbl_Non_Digit_Chars	LABELER	Labels all non-digit characters consistently.
lbl_nzl_bank_acct_number	LABELER	Labels New Zealand bank account numbers.

Name	Asset Type	Description
lbl_nzl_phone	LABELER	Labels New Zealand phone numbers.
lbl_nzl_social_welfare_no	LABELER	Labels a New Zealand Social Welfare Number using a regular expression. No checksum validation as information on checksum is not available.
lbl_word_symbol	LABELER	Labeler identifying Unicode definition of a word combined with one or more symbols. For example: email addresses
mplt_aus_address_validation_hybrid	DMAPPLET	Validates the deliverability of Australian addresses. Use the mapplet when you can connect the input address fields to fields on the Multiline model.
mplt_aus_bank_account_number_parse	DMAPPLET	Parses a 7-digit to 13-digit string as an Australian bank account number from a longer string.
mplt_aus_bank_account_number_standardize	DMAPPLET	Standardizes a 7-digit to 13-digit string as an Australian bank account number, regardless of spacing or punctuation. If the input value is not between 7 and 13 digits in length, the mapplet returns the input value. The mapplet returns the numbers without punctuation.
mplt_aus_bank_account_number_validate	DMAPPLET	Validates an Australian bank account number between 7-digit to 13-digit. Checks the format and length of the Australian bank account number. The mapplet parses strings of numbers in the following formats: - nnnnnnn - r nnnnnnnnn - nnnnnnnnnnnn
mplt_aus_bsb_parse	DMAPPLET	Parses a 6-digit string as an Australian BSB (Bank-State-Branch) number from a longer string. The mapplet parses strings of numbers in the following format: - 111111
mplt_aus_bsb_standardize	DMAPPLET	Standardizes Australian BSB (Bank-State-Branch) number as indicated by the variable value in the expression. The mapplet returns strings of numbers in the following format: - nnn-nnn If any alphanumeric character or an input value is not equal to 6 digits, the mapplet returns the input value.

Name	Asset Type	Description
mplt_aus_bsb_validate	DMAPPLET	Validates a 6-digit string as an Australian BSB (Bank-State-Branch) number after checking the length and the format of the number. The mapplet validates strings of numbers in the following format: - 111-111
mplt_aus_business_number_parse	DMAPPLET	Parses 11-digit numeric strings as Australian Business Numbers (ABN) from a string of text.
mplt_aus_business_number_standardize	DMAPPLET	Standardizes a 11-digit Australian Business Numbers (ABN) to NN NNN NNN NNN, regardless of spacing or punctuation. If any alphanumeric character or an input value is less than 11 digit, the mapplet returns the input value.
mplt_aus_business_number_validate	DMAPPLET	Validates Australian Business numbers (ABN). Rule assumes that the ABN is formatted by mplt_AUS_Business_Number_Standardize or is an 11 digit numeric string.
mplt_aus_company_name_standardize	DMAPPLET	Standardizes a company name and provides the acronym if possible.
mplt_aus_company_number_parse	DMAPPLET	Parses an Australian Company Number (ACN), which is a unique 9-digit identifier assigned by the Australian Securities and Investments Commission (ASIC).
mplt_aus_company_number_standardize	DMAPPLET	Standardizes the format of an input Australian Company Number (ACN), which is a unique 9-digit identifier assigned by the Australian Securities and Investments Commission (ASIC).
mplt_aus_company_number_validate	DMAPPLET	Validates an Australian Company Number (ACN), which is 9-digit number when registered. The Australian Company Number (ACN) must be displayed on all company documents.
mplt_aus_covid_document_number_parse	DMAPPLET	Parses a number that fits the COVID document number format.
mplt_aus_covid_document_number_standardize	DMAPPLET	Standardizes the input to the COVID document number format, the default being nnnn-nnnn-nnnn. The option to standardize with a space or no punctuation is available by updating the *Covid_number_Format* variable in the expression asset.
mplt_aus_covid_document_number_validate	DMAPPLET	Validates that the input represents a COVID document number.
mplt_aus_crn_parse	DMAPPLET	Parses an Australian Centrelink Reference Number (CRN) from a string of text. The CRN consists of nine digits and an alphabet at the end.

Name	Asset Type	Description
mplt_aus_crn_standardize	DMAPPLET	Standardizes Australian Centrelink Reference Number (CRN) as indicated by the variable value in the expression. The mapplet returns strings of numbers in the following formats: - No_Punctuation. `nnnnnnnnn` - Space. `nnn nnn nnn` Default is No_Punctuation.
mplt_aus_crn_validate	DMAPPLET	Validates Australian Centrelink Reference Number (CRN). Returns Valid or Invalid.
mplt_aus_family_name_and_address_match	DMAPPLET	Identify duplicate rows in Australian data based on personal surname and address data grouped by the first two digits of the postcode and the surname.
mplt_aus_gender_assignment	DMAPPLET	Assigns gender according to first names. The mapplet returns "M" for male names, "F" for female names, and "U" if the gender is unknown. For example, the mapplet assigns the name "John Smith" a gender of "M" for male.
mplt_aus_given_name_standardize	DMAPPLET	Generates given names from Australian nicknames. For example, the mapplet standardizes the nickname "Bob" to the given name "Robert."
mplt_aus_healthcare_identifier_parse	DMAPPLET	Parses a 16-digit string as an Australian Health Identifier from a longer string. The Health Identifier contains the following identifiers: - IHI. Individual Healthcare Identifier, beginning with 800360. - HPI-I. Healthcare Provider Identifier, beginning with 800361. - HPI-O. Healthcare Provider Identifier, beginning with 800362. The mapplet parses strings of numbers in the following format: - `800360/1/2 NNNNNNNNNN`
mplt_aus_healthcare_identifier_standardize	DMAPPLET	Standardizes Australian Healthcare Identifiers as indicated by the variable value in the expression. The Health Identifier contains the following identifiers: - IHI. Individual Healthcare Identifier. - HPI-I. Healthcare Provider Identifier. - HPI-O. Healthcare Provider Identifier. The mapplet returns strings of numbers in the following formats: - No_Punctuation. `nnnnnnnnnnnnnnn` - Space. `nnnn nnnn nnnn nnnn` - Dash. `nnnn-nnnn-nnnn-nnnn`

Name	Asset Type	Description
mplt_aus_healthcare_identifier_validate	DMAPPLET	Validates a 16-digit string as an Australian Health Identifier after checking the length and the format. The Health Identifier contains the following identifiers: - IHI. Individual Healthcare Identifier. - HPI-I. Healthcare Provider Identifier. - HPI-O. Healthcare Provider Identifier. The mapplet validates numbers in the following formats: - `1111-1111-1111-1111` - `1111 1111 1111 1111` - `1111111111111111`
mplt_aus_individual_name_and_address_match	DMAPPLET	Identifies duplicate rows based on person names and Australian address data. The mapplet generates group keys from the name data.
mplt_aus_individual_name_and_email_match	DMAPPLET	Identifies duplicate rows based on email addresses and Australian person names. The mapplet generates group keys from the email address data.
mplt_aus_individual_name_and_tfn_match	DMAPPLET	Identifies duplicate rows for Australian data based on Tax File Numbers (TFN) and person names. The mapplet generates group keys from the TFN data.
mplt_aus_individual_name_date_match	DMAPPLET	Identifies duplicate rows based on Australian person names and dates. The mapplet generates group keys from the date data.
mplt_aus_individual_name_match	DMAPPLET	Identifies duplicate rows based on Australian person names. The mapplet generates group keys from the surname values.
mplt_aus_individual_name_phone_match	DMAPPLET	Identifies duplicate rows based on Australian person names and telephone numbers. The mapplet generates group keys from the telephone number data.
mplt_aus_medicare_number_parse	DMAPPLET	Parses an Australian Medicare number from a string of text into its own field. For example, parses `2428778` from `test2428 77813 2-1test`.
mplt_aus_medicare_number_standardize	DMAPPLET	Standardizes a 10-digit or 11-digit string as an Australian Medicare number, regardless of spacing or punctuation. If input value is not equal to 11-digit, the mapplet returns the input value. The mapplet returns strings of numbers in the following format: - `nnnn nnnnn n`

Name	Asset Type	Description
mplt_aus_medicare_number_validate	DMAPPLET	Validates a 11-digit number as an Australian Medicare number. Checks the format and length of Australian Medicare number. Checksum validation included.
mplt_aus_personal_name_and_data_match	DMAPPLET	Identifies duplicate rows in Australian data based on person names and personal data. The fields in the personal data column must contain a single type of data, such as a telephone number, email, or Tax File Number. The mapplet generates group keys from the personal data.
mplt_aus_phone_number_parse	DMAPPLET	Parses an Australian telephone number from a string. For example, parses `642123456` from test `642123456test`.
mplt_aus_phone_number_standardize	DMAPPLET	Standardizes the format of an Australian phone number country code, area code and landline number to the following formats: - Standard. `+nn (nn) nnnnn` - Dashes. `+nn-nnnn-nnnn` - No spaces. `+nnnnnnnnnn`
mplt_aus_phone_number_validate	DMAPPLET	Validates the area code and length of Australian telephone numbers. The mapplet returns the region of the telephone number as well as codes that indicate if the area code and length of a telephone number is valid.
mplt_aus_tax_file_number_parse	DMAPPLET	Parses a 9-digit string as an Australian Tax File Number (TFN) from a string of text, for example `123456789` or `113-456-789`
mplt_aus_tax_file_number_standardize	DMAPPLET	Standardizes Australian Tax File Numbers (TFN). The mapplet returns strings of numbers in the following formats: - No_Punctuation. `nnnnnnnnn` - Space. `nnn nnn nnn` - Dash. `nnn-nnn-nnn`
mplt_aus_tax_file_number_validate	DMAPPLET	Validates Australia Tax File Number (TFN). For example, `289898765` or `234-345-654`
mplt_AV6_Assign_DQ_Status_Code_Description	DMAPPLET	Describes the Match Score value that a verifier asset returns for the Address Verification version 6 engine.
mplt_Luhn_Algorithm	DMAPPLET	Applies the Luhn algorithm, a mod 10 calculation commonly used for credit card number check digit validation, to an input value. Returns Valid if the supplied check digit matches the calculated value. Returns Invalid if the supplied check digit does not match the calculated value.

Name	Asset Type	Description
mplt_nzl_bank_account_number_standardize	DMAPPLET	Standardizes New Zealand bank account number. The domestic account numbers include the 4-digit branch number, 7-digit account, and 3-digit suffix, in the following format: `NN-NNNN-NNNNNNN-NNN`.
mplt_nzl_bank_account_number_validate	DMAPPLET	Parses a 14-digit to 16-digit string as a New Zealand Bank account number or BSB number from a longer string. Checks the format and length of New Zealand Bank account number or BSB number. The mapplet parses strings of numbers in the following format: - `882896855917503` - `23 2042 5076936 050` - `04-49313241531-00`
mplt_nzl_ird_number_parse	DMAPPLET	Parses 9-digit numeric strings as New Zealand Inland Revenue Department numbers (IRD) from a string of text.
mplt_nzl_ird_number_standardize	DMAPPLET	Standardizes New Zealand Inland Revenue Dept. number (IRD) as indicated by the variable value in the expression. The mapplet parses strings of numbers in the following format: - No_Punctuation. `nnnnnnnnn` - Space.`nnn nnn nnn` - Dash. `nnn-nnn-nnn` Default is No_Punctuation.
mplt_nzl_ird_number_validate	DMAPPLET	Validates New Zealand Inland Revenue Department numbers (IRD) number. Returns Valid or Invalid.
mplt_nzl_phone_number_parse	DMAPPLET	Parses a New Zealand telephone number from a string. For example, parses `642123456` from test `642123456test.`
mplt_nzl_phone_number_standardize	DMAPPLET	Standardizes the format of a New Zealand phone number country code, area code and landline number to the following formats: - Standard. `+nn (nnn) nnnnn` - Dashes. `+nn-nnnn-nnnn` - No spaces. `+nnnnnnnnnn`
mplt_nzl_phone_number_validate	DMAPPLET	Validates the area code and length of New Zealand telephone numbers. The mapplet returns the region of the telephone number as well as codes that indicate if the area code and length of a telephone number is valid.
mplt_nzl_social_welfare_number_validate	DMAPPLET	Validates a New Zealand Social Welfare Number from a string with limited punctuation. Checks the length and format of the number. No checksum validation as information on checksum is not available.

Name	Asset Type	Description
mplt_Remove_Non_Numbers	DMAPPLET	Removes all characters that are not a digit or number from an input string.
mplt_Remove_Punctuation	DMAPPLET	Removes all punctuation symbols from the input.
mplt_Remove_Punctuation_and_Space	DMAPPLET	Removes all punctuation symbols and character spaces from an input field, and returns the remaining digits and letters.
mplt_Replace_Limited_Punct_with_Space	DMAPPLET	Replaces punctuation symbols including a forward slash, backslash, exclamation point, period, or underscore character with a character space. Also replaces instances of two, three, or four consecutive character spaces with a single space.
mplt_UpperCase	DMAPPLET	Changes characters to uppercase.
nzl_tel_area_codes_infa	DICTIONARY	Contains a list of New Zealand telephone area codes.
p_assign_prename_gender	PARSE	Assigns gender depending on honorifics.
p_aus_area_code	PARSE	Parses an Australian area code.
p_aus_assign_midname_gender	PARSE	Parses the gender for the assigned middle name.
p_aus_bank_acct	PARSE	Parses an Australian Bank account number.
p_aus_bsb_number	PARSE	Parses an Australian BSB (Bank-State-Branch) Number.
p_aus_company_number	PARSE	Parses an Australian Company Number (ACN) from a string.
p_aus_crn	PARSE	Parses an Australian Customer Reference Number (CRN), which consists of nine digits and an alphabet at the end.
p_aus_health_id	PARSE	Parses a 16-digit IHI number.
p_aus_int_country_code	PARSE	Parses the Australia international dialing country codes.
p_aus_medicare_no	PARSE	Parses an Australian Medicare number from a string.
p_aus_phone	PARSE	Parses an Australian phone number.
p_nzl_area_codes	PARSE	Parses specific telephone codes.
p_nzl_int_country_code	PARSE	Parses a New Zealand country code.
p_nzl_phone	PARSE	Parses a New Zealand phone number with the phone code.

Name	Asset Type	Description
p_nzl_phone_number	PARSE	Parses a New Zealand phone number.
p_nzl_phone_parse	PARSE	Parses a New Zealand phone number from a string.
p_parse_abn	PARSE	Parses an Australian Business number from a string of text.
p_parse_aus_covid_number	PARSE	Parses an Australian COVID number.
p_parse_aus_tfn	PARSE	Parses an Australian tax file number (TFN).
p_parse_full_name	PARSE	Parses using the pre-built Australia Multiname parsing.
p_parse_nzl_ird	PARSE	Parses an 8-digit and 9-digit New Zealand Inland Revenue Dept. (IRD) number from a string of text.
v_aus_address_validation_hybrid	VERIFIER	Verifies Australian addresses with input fields that map to the Hybrid model.

CHAPTER 8

Data Quality bundle for Europe

Use the Data Quality bundle for Europe to accelerate the configuration and deployment of data quality solutions in your organizations. The bundle includes mapplets and other assets that define cleansing, labeling, and parsing operations for the data of these countries.

The Data Quality bundle for Europe includes assets for data in the following countries:

- France
- Germany
- Italy
- Portugal
- Spain

The following table lists the assets contained in the Data Quality bundle for Europe:

Name	Country	Asset Type	Description
c_fra_get_name_standard	France	Cleanse	Replaces the input nicknames with a standard first name. For example, the asset standardizes the nickname "Beth" to "Elizabeth."
c_fra_remove_intl_dialcodes	France	Cleanse	Removes the French international dial codes.
c_fra_salutation_assignment	France	Cleanse	Removes custom punctuation marks from an input.
c_merge_fields	France	Cleanse	Merges the Overflow, Before, After and Unparsed fields from the parse asset.
fra_nicknames_infa	France	Dictionary	Contains a list of French nicknames.
fra_phone_area_code_infa	France	Dictionary	Contains a list of French telephone area codes.
lbl_fra_cni	France	Labeler	Labels French Carte Nationale d'Identité (CNI) numbers as valid.
lbl_INSEE_Format	France	Labeler	Labels French Institut National de la Statistique et des Études Économiques (INSEE) numbers as Valid.

Name	Country	Asset Type	Description
lbl_fra_passport	France	Labeler	Labels French passport numbers as Valid.
mplt_FRA_CNI_Number_Standardize	France	Mapplet	Standardizes a 9-digit and 12-digit character as a French Carte Nationale d'Identité (CNI) number, regardless of spacing or punctuation.
mplt_FRA_CNI_Number_Validation	France	Mapplet	Validates French Carte Nationale d'Identité (CNI) number.
mplt_FRA_CNI_Number_Parse	France	Mapplet	Parses a French Carte Nationale d'Identité (CNI) number from a string of text. The CNI consists of 12-digit or 9-digit alphanumeric character. For example, `123456789ABC` or `113ABC123`.
mplt_FRA_Driving_License_Parse	France	Mapplet	Parses a 12-digit French driving license number from a string of text.
mplt_FRA_Driving_License_Standardize	France	Mapplet	Standardizes a 12-digit French driving license number to `NNNNNNNNNNNN` format, regardless of spacing or punctuation. If input value is not equal to 12-digit, the mapplet returns the input value.
mplt_FRA_Driving_License_Validation	France	Mapplet	Validates French driving license number. Checks the length of French driving license number. Returns Valid or Invalid.
mplt_FRA_Given_Name_Standardize	France	Mapplet	Generates given names from French nicknames. For example, the mapplet standardizes the nickname "Bob" to the given name "Robert."
mplt_FRA_INSEE_Number_Standardize	France	Mapplet	Standardizes the French Institut National de la Statistique et des Études Économiques (INSEE) number.
mplt_FRA_INSEE_Number_Validation	France	Mapplet	Validates the French Institut National de la Statistique et des Études Économiques (INSEE) number.
mplt_FRA_INSEE_Number_Parse	France	Mapplet	Parses the French Institut National de la Statistique et des Études Économiques (INSEE) number from a string of text.
mplt_FRA_Passport_Number_Parse	France	Mapplet	Parses a French passport number from a string of text.

Name	Country	Asset Type	Description
mplt_FRA_Passport_Number_Standardize	France	Mapplet	Standardizes a 9-digit character as a French passport number, regardless of spacing or punctuation.
mplt_FRA_Passport_Number_Validation	France	Mapplet	Validates a French passport number.
mplt_FRA_Phone_Number_Parse	France	Mapplet	Parses a French telephone number from a string of text.
mplt_FRA_Phone_Number_Validation	France	Mapplet	Validates a French telephone number.
mplt_FRA_Phone_Number_Standardize	France	Mapplet	Standardizes French telephone numbers to international and local dialing formats. The mapplet standardizes any telephone number that can be handled by the French telephone parsing rule.
mplt_FRA_Salutation_Assignment	France	Mapplet	Generates formal and casual greetings from prefixes and names. For example, when input data contains "Mr. John Smith," the mapplet generates the formal greeting "Cher Mr. Smith," and the casual greeting "Cher John,'. **Note:** The Gender input should be a single character, M or F.
mplt_FRA_SIREN_Number_Parse	France	Mapplet	Parses a French Système d'identification du répertoire des entreprises (SIREN) number from a string of text.
mplt_FRA_SIREN_Number_Standardize	France	Mapplet	Standardizes French Système d'identification du répertoire des entreprises (SIREN) numbers.
mplt_FRA_SIREN_Number_Validation	France	Mapplet	Validates French Système d'identification du répertoire des entreprises (SIREN) numbers.
mplt_FRA_SIRET_Number_Parse	France	Mapplet	Parses the French système d'identification du répertoire des établissements (SIRET) number from a string of text.
mplt_FRA_SIRET_Number_Standardize	France	Mapplet	Standardizes a French système d'identification du répertoire des établissements (SIRET) number.
mplt_FRA_SIRET_Number_Validation	France	Mapplet	Validates French système d'identification du répertoire des établissements (SIRET) numbers.
p_fra_area_code	France	Parse	Parses a French area code from the input data.

Name	Country	Asset Type	Description
p_fra_dl	France	Parse	Parses a French driver's license number from a string of text.
p_fra_phone_number	France	Parse	Parses a French telephone number from the input data.
p_siren_parse	France	Parse	Parses a French Système d'identification du répertoire des entreprises (SIREN) number from the input data.
p_fra_siret	France	Parse	Parses a French système d'identification du répertoire des établissements (SIRET) number from the input data.
p_fra_passport	France	Parse	Parses a France passport number from the input data.
p_fra_cni_parse	France	Parse	Parses a French Carte Nationale d'Identité (CNI) number from the input data.
p_fra_insee_parse	France	Parse	Parses the French Institut National de la Statistique et des Études Économiques (INSEE) number from the input data.
p_fra_phone_number	France	Parse	Parses a French telephone number from the input data.
c_DEU_tel_areacode_region	Germany	Cleanse	Returns `valid` for correct German telephone area codes.
c_remove_intl_dialling_code	Germany	Cleanse	Standardizes the international dialing code to `49`.
c_remove_non_numbers	Germany	Cleanse	Removes non-number from a string.
c_remove_symbols_chars	Germany	Cleanse	Removes symbols and letters.
deu_tel_area_codes_infa	Germany	Dictionary	Contains a list of German telephone area codes.
lbl_deu_dl	Germany	Labeler	Labels a valid German driving license number as valid.
lbl_deu_id_card_number	Germany	Labeler	Labels a valid German ID card number as valid.
lbl_deu_passport	Germany	Labeler	Labels a valid German passport number as valid based on the format.
lbl_deu_vat	Germany	Labeler	Labels a German VAT (Value Added Tax) number as valid.

Name	Country	Asset Type	Description
lbl_word_symbol	Germany	Labeler	Labels unicode definition of a word combined with one or more symbols. For example: email addresses
mplt_DEU_Driver_License_Number_Parse	Germany	Mapplet	Parses an eleven-character alphanumeric number as a German driving license number from a longer string.
mplt_DEU_Driver_License_Number_Standardize	Germany	Mapplet	Standardizes an eleven-character German driving license number, regardless of spacing or punctuation. If input value is not equal to eleven-characters, the mapplet returns the input value. The mapplet returns the driving license number in the following format: `Y66EPG3020F`
mplt_DEU_Driver_License_Number_Validation	Germany	Mapplet	Validates if an eleven-character alphanumeric number corresponds to a German driving license number.
mplt_DEU_ID_Card_Number_Parse	Germany	Mapplet	Parses a 10-digit or 11-digit alphanumeric code as a German ID card number from a longer string. The German ID card number is also known as the Personalausweisnummer. A German ID card number is a unique identification number assigned to each individual who holds a German national identity card (Personalausweis).
mplt_DEU_ID_Card_Number_Standardize	Germany	Mapplet	Standardizes a 9-digit to 11-digit character as a German ID card number, regardless of spacing or punctuation.
mplt_DEU_ID_Card_Number_Validation	Germany	Mapplet	Validates a 9-digit, 10-digit, or 11-digit alphanumeric code as a German ID card number. The German ID card number is a unique identification number assigned to each individual who holds a German national identity card (Personalausweis).
mplt_DEU_Passport_Number_Parse	Germany	Mapplet	Parses a 9-digit or 11-digit alphanumeric number as a German passport number from a longer string.

Name	Country	Asset Type	Description
mplt_DEU_Passport_Number_Standardize	Germany	Mapplet	Standardizes a 9-digit to 11-digit character as a German passport number, regardless of spacing or punctuation. The mapplet returns strings of numbers in the following format: - nnnnnnnnn - nnnnnnnnnn - nnnnnnnnnnn If input value is not equal to 9-digit, 10-digit, or 11-digit, and an input value is less than 8 characters or greater than 11 digit, the mapplet returns the input value.
mplt_DEU_Passport_Number_Validation	Germany	Mapplet	Validates a German passport number.
mplt_DEU_Phone_Number_Parse	Germany	Mapplet	Parses a German telephone number from a longer string. The phone number can be prefixed with the international dialing information or can be in local format.
mplt_DEU_Phone_Number_Standardization	Germany	Mapplet	Standardizes German telephone numbers to international and local dialing formats. The mapplet matches to known area codes and format accordingly. Phone numbers that do not match a known area code are passed through.
mplt_DEU_Phone_Number_Validation	Germany	Mapplet	Validates Germany telephone numbers.
mplt_DEU_tax_file_number_parse	Germany	Mapplet	Parses a German Tax File Number (TFN) from a string of text.
mplt_DEU_Tax_File_Number_Standardize	Germany	Mapplet	Standardizes German Tax File Numbers (TFN).
mplt_DEU_Tax_File_Number_Validation	Germany	Mapplet	Validates Germany Tax File Number (TFN).
mplt_DEU_VAT_Number_Parse	Germany	Mapplet	Parses a German VAT (Value Added Tax) number from a longer string. The German VAT number is a 9-digit number or DE followed by 9-digits. The mapplet parses strings of numbers in the following formats: - 999999999

Name	Country	Asset Type	Description
mplt_DEU_VAT_Number_Standardize	Germany	Mapplet	Standardizes a 11-digit alphanumeric character as a German VAT (Value Added Tax) number, regardless of spacing or punctuation. The mapplet returns strings of numbers in the following format: `DE581901216` If input value is not equal to 11-digit and does not match the pattern, the mapplet returns the input value.
mplt_DEU_VAT_Number_Validation	Germany	Mapplet	Validates a German VAT (Value Added Tax) number.
p_DEU_country_code	Germany	Parse	Parses German country codes from the input data.
p_deu_dl	Germany	Parse	Parses a German driving license number from a longer string.
p_deu_id	Germany	Parse	Parses a German ID card number.
p_deu_passport	Germany	Parse	Parses a German passport number.
p_deu_phone_number	Germany	Parse	Parses German phone number from input data.
p_deu_vat	Germany	Parse	Parses a German VAT (Value Added Tax) number from a longer string.
p_deu_tfn	Germany	Parse	Parses German Tax File Numbers (TFN) from input data.
p_match_area_code_2_digit	Germany	Parse	Parses two-digit German telephone area codes.
p_match_area_code_3_digit	Germany	Parse	Parses three-digit German telephone area codes.
p_match_area_code_4_digit	Germany	Parse	Parses four-digit German telephone area codes.
p_match_area_code_5_digit	Germany	Parse	Parses five-digit German telephone area codes.
p_match_area_code_6_digits	Germany	Parse	Parses six-digit German telephone area codes.
ita_commune	Italy	Dictionary	Contains a list of Italian comune or municipalities.
c_ITA_Commune	Italy	Cleanse	Validates an Italian commune.
mplt_ITA_Codice_Fiscale_Parse	Italy	Mapplet	Parses an Italian fiscal code or Codice fiscale from a string of text.

Name	Country	Asset Type	Description
mplt_ITA_Comune_Validation	Italy	Mapplet	Validates a comune name in Italy. A comune is a basic administrative division in Italy, roughly equivalent to a township or municipality.
mplt_ITA_FiscalCode_Validation	Italy	Mapplet	Validates a 16-digit alphanumeric code as an Italian fiscal code based on the determined format.
mplt_ITA_Phone_Number_Parse	Italy	Mapplet	Parses an Italian telephone number from a longer string. The phone number can be prefixed with the international dialing information or can be in local format.
mplt_ITA_Phone_Number_Standardization	Italy	Mapplet	Standardizes Italian telephone numbers to international and local dialing formats. The mapplet checks and matches a known area code. Phone numbers that do not match a known area code are passed through.
mplt_ITA_Phone_Number_Validation	Italy	Mapplet	Validates Italian telephone numbers.
mplt_ITA_Postcode_Parse	Italy	Mapplet	Parses Italian postal codes using a dictionary.
mplt_ITA_Postcode_Standardize	Italy	Mapplet	Standardizes Italian postal codes. The mapplet returns strings of numbers in the following format: `nnnnn` If any input value is not equal to 5 digits, the mapplet returns the input value.
mplt_ITA_Registration_Number_Plate_Validation	Italy	Mapplet	Validates Italian registration number plates.
mplt_ITA_VAT_Number_Validation	Italy	Mapplet	Validates an Italian VAT (Value Added Tax) number using a regular expression and Luhn algorithm, a mod 10 calculation for check digit validation.
p_ita_fiscale_code	Italy	Parse	Parses an Italian fiscale code from the input data.
p_ita_area_codes	Italy	Parse	Parses Italian telephone area codes.
p_ita_phone	Italy	Parse	Parses an Italian telephone number from an input string.
p_ita_postcodes	Italy	Parse	Parses Italian postal codes.
c_ita_label_alphabets	Italy	Cleanse	Labels the alphabets as X.

Name	Country	Asset Type	Description
c_ita_label_number	Italy	Cleanse	Labels the numbers as 9.
c_ita_remove_alphabet	Italy	Cleanse	Removes alphabets from an input value.
c_ita_remove_labels	Italy	Cleanse	Removes X from a string.
c_ita_strip_prefix	Italy	Cleanse	Strips Italian telephone prefix from the start of the input string.
c_remove_non_number_characters	Italy	Cleanse	Removes characters that are not digit or number from a string.
ita_postcodes_infa	Italy	Dictionary	Contains a list of Italian postal codes
ita_tel_area_codes_infa	Italy	Dictionary	Contains a list of Italian telephone area codes.
lbl_ita_ifc	Italy	Labeler	Validates an Italian fiscal code.
lbl_ita_registration_number	Italy	Labeler	Checks the regular expression pattern and labels the valid registration number plate values.
lbl_ita_vat	Italy	Labeler	Labels valid Italian VAT (Value Added Tax) numbers.
c_merge_fields	Portugal	Cleanse	Merges the Overflow, Before, After and Unparsed fields from the parse asset.
c_PRT_remove_intl_dialing_code	Portugal	Cleanse	Removes the Portuguese international dialing telephone codes.
c_PRT_tel_areacode_region	Portugal	Cleanse	Returns `valid` for correct Portugal telephone area codes.
prt_tel_area_codes_infa	Portugal	Dictionary	Contains a list of Portugal telephone area codes.
mplt_PRT_NIF_Parse	Portugal	Mapplet	Parses a Número de Identificação Fiscal (NIF) number from a string of text.
mplt_PRT_NIF_Standardize	Portugal	Mapplet	Standardizes the Número de Identificação Fiscal (NIF) number.
mplt_PRT_NIF_Validation	Portugal	Mapplet	Validates the Número de Identificação Fiscal (NIF) number.
mplt_PRT_NIPC_Parse	Portugal	Mapplet	Parses a Número de Identificação Pessoa Colectiva (NIPC) number from a string of text.
mplt_PRT_NIPC_Standardize	Portugal	Mapplet	Standardizes a Número de Identificação Pessoa Colectiva (NIPC) number.

Name	Country	Asset Type	Description
mplt_PRT_NIPC_Validation	Portugal	Mapplet	Validates a Número de Identificação Pessoa Colectiva (NIPC) number.
mplt_PRT_Phone_Number_Parse	Portugal	Mapplet	Parses a Portuguese telephone number. The mapplet uses p_PRT_phone_number. The mapplet also merges the additional text to an output field.
mplt_PRT_Phone_Number_Standardization	Portugal	Mapplet	Standardizes Portugal telephone numbers.
mplt_PRT_Phone_Number_Validation	Portugal	Mapplet	Validates Portugal telephone numbers.
p_PRT_phone_number	Portugal	Parse	Extracts a Portuguese telephone number from the input data, and merges the unparsed text.
p_PRT_NIF_number	Portugal	Parse	Parses the Número de Identificação Fiscal (NIF) number from the input data.
p_PRT_NIPC_number	Portugal	Parse	Parses the Número de Identificação Pessoa Colectiva (NIPC) number from the input data.
p_PRT_phone_number_old	Portugal	Parse	Extracts a Portuguese telephone number from the input data, and merges the unparsed text.
c_esp_get_destination_code	Spain	Cleanse	Standardizes a list of Spanish destination codes using dictionary.
c_esp_nie_std	Spain	Cleanse	Removes custom strings.
c_esp_remove_dialing_codes	Spain	Cleanse	Removes the Spanish international dialing codes from an input string.
c_esp_remove_intl_dialing_code	Spain	Cleanse	Removes the Spanish international dialing telephone codes.
c_merge_fields	Spain	Cleanse	Merges the Overflow, Before, After and Unparsed fields from the parse asset.
c_nif_type_check	Spain	Cleanse	Identifies the type of Certificado de Identificación Fiscal (CIF) number being entered. For example, a Corporation, LLC, Public Agency, or Foreign Entity.
esp_cif_type	Spain	Dictionary	Indicates the type of the tax or Certificado de Identificación Fiscal (CIF) entity.
esp_city_from_phone_code_infa	Spain	Dictionary	Indicates the city from Spain city telephone codes.

Name	Country	Asset Type	Description
lbl_esp_dl	Spain	Labeler	Validates a driving license number for Spain.
lbl_ESP_passport	Spain	Labeler	Validates Spanish passport numbers.
lbl_esp_repeating_numbers	Spain	Labeler	Labels repeating numbers that contain 7-digits. For example, labels the repeating number `5555555`.
lbl_Repeating_Digits	Spain	Labeler	Labels repeating numbers that contains different patterns. For example, labels repeating patterns 202020, 898989, or 454545.
lbl_repeating_numbers	Spain	Labeler	Labels repeating numbers that contain three or more digits. For example, labels the repeating number `555`.
lbl_repeating_numbers_nie	Spain	Labeler	Labels repeating Spanish Número de Identidad de Extranjero (NIE) numbers that contain three or more digits. For example, labels the repeating number `555`.
lbl_repeating_numbers_phone	Spain	Labeler	Labels repeating numbers that contain three or more digits in a telephone number. For example, labels the repeating numbers from 000 to 999.
mplt_ESP_CIF_Parse	Spain	Mapplet	Parses the Spanish Certificado de Identificación Fiscal (CIF) number from a string.
mplt_ESP_CIF_Standardization	Spain	Mapplet	Standardizes the Spanish Certificado de Identificación Fiscal (CIF). The CIF is the tax ID number that identifies a company or a legal entity in Spain.
mplt_ESP_CIF_Validation	Spain	Mapplet	Validates the Spanish Certificado de Identificación Fiscal (CIF). The CIF is the tax ID number that identifies a company or a legal entity in Spain.
mplt_ESP_Driving_License_Parse	Spain	Mapplet	Parses a Spanish driving license number from a larger string.
mplt_ESP_Driving_License_Validation	Spain	Mapplet	Validates Spanish driving license numbers.
mplt_esp_dni_parse	Spain	Mapplet	Parses the Spanish Documento Nacional de Identidad (DNI), which consists of eight digits followed by a character, in the following format: `NNNNNNNNC`.

Name	Country	Asset Type	Description
mplt_ESP_DNI_Validation	Spain	Mapplet	Validates the Spanish Documento Nacional de Identidad (DNI). Returns Valid, Invalid, and Suspect. The Suspect value has a valid check digit but contains repeated digits. For example: `11111111H`.
mplt_esp_dni_standardize	Spain	Mapplet	Standardizes the Spanish Documento Nacional de Identidad (DNI).
mplt_ESP_Driving_License_Standardize	Spain	Mapplet	Standardizes a Spanish driving license number.
mplt_ESP_NIE_Parse	Spain	Mapplet	Parses the Spanish Número de Identidad de Extranjero (NIE) from a larger string. The number is issued to all foreigners once they obtain residency, and the number appears on their Residency Card.
mplt_ESP_NIE_Standardization	Spain	Mapplet	Standardizes the Spanish Número de Identidad de Extranjero (NIE). For non Spanish people legally living in Spain, any letters, symbols, or spaces are removed.
mplt_ESP_NIE_Validation	Spain	Mapplet	Validates the Spanish Número de Identidad de Extranjero (NIE). Returns Valid, Invalid, and Suspect. The Suspect value has a valid check digit but contains repeated digits. For example: `X1111111G`
mplt_ESP_NIF_Parse	Spain	Mapplet	Parses the Spanish Número de Identificación Fiscal (NIF) from a larger string.
mplt_ESP_NIF_Standardization	Spain	Mapplet	Standardizes the Spanish Número de Identificación Fiscal (NIF).
mplt_ESP_NIF_Validation	Spain	Mapplet	Validates the Spanish Número de Identificação Fiscal (NIF).
mplt_ESP_Passport_Number_Parse	Spain	Mapplet	Parses a Spanish passport number from a string of text.
mplt_ESP_Passport_Number_Standardize	Spain	Mapplet	Standardizes Spanish passport numbers.
mplt_ESP_Passport_Number_Validation	Spain	Mapplet	Validates Spanish passport numbers.

Name	Country	Asset Type	Description
mplt_ESP_Phone_Number_Standardization	Spain	Mapplet	Standardizes Spanish telephone numbers to international and local dialing formats.
mplt_ESP_Phone_Number_Validation	Spain	Mapplet	Validates Spanish telephone numbers.
mplt_ESP_Phone_Parse	Spain	Mapplet	Parses a Spanish telephone number from a larger string.
mplt_ESP_Phone_Parse_Standardize_Validate	Spain	Mapplet	Parses a Spanish telephone number, validates and standardizes it to the International and local formats. Returns status codes of Valid, Invalid, and Suspect. Suspect numbers have repeating digits.
mplt_ESP_Phone_Number_Validation	Spain	Mapplet	This mapplet checks if the number is a valid Spain phone number. Tags: PII, Europe, Validate, Spain, Phone
mplt_ESP_Phone_Parse_Standardize_Validate	Spain	Mapplet	Rule parses a Spanish phone number, validates and standardizes it to the International and local formats. Returns status codes of Valid, Invalid and Suspect. Suspect numbers have repeating digits. Tags: Spain, Phone, High, PII,
p_esp_nif	Spain	Parse	Parses the Spanish Número de Identificación Fiscal (NIF) number from the input data.
p_esp_cif	Spain	Parse	Parses the Spanish Certificado de Identificación Fiscal (CIF) number from a longer string.
p_esp_dni	Spain	Parse	Parses the Spanish Documento Nacional de Identidad (DNI), which consists of eight digits followed by a character.
p_esp_nie	Spain	Parse	Parses the Spanish Número de Identidad de Extranjero (NIE) number.
p_esp_phone	Spain	Parse	Parses a Spanish phone number from an input string.
p_esp_dl	Spain	Parse	Parses a Spanish driving license number from an input string.
p_esp_passport_parse	Spain	Parse	Parses a Spanish passport number from an input data.

Name	Country	Asset Type	Description
mplt_Remove_Punctuation_and_Space	Core Bundle	Mapplet	Removes all punctuation symbols and character spaces from an input field, and returns the remaining digits and letters.
mplt_Remove_Non_Numbers	Core Bundle	Mapplet	Removes all characters that are not a digit or number from an input string.
mplt_UpperCase	Core Bundle	Mapplet	Changes alphabetic characters to uppercase.
mplt_Luhn_Algorithm	Core Bundle	Mapplet	Applies the Luhn algorithm, a mod 10 calculation commonly used for credit card number check digit validation, to an input value. Returns Valid if the supplied check digit matches the calculated value. Returns Invalid if the supplied check digit does not match the calculated value.
mplt_Remove_Hyphen	Core Bundle	Mapplet	Removes all hyphens from an input string.
mplt_Remove_Leading_Zero	Core Bundle	Mapplet	Removes a leading zero from a string.
mplt_Remove_Period_Parantheses	Core Bundle	Mapplet	Removes all occurrences of a period and left and right parentheses in a string.
mplt_Remove_Space	Core Bundle	Mapplet	Removes all occurrences of a character space from an input field.
mplt_Replace_Limited_Punct_with_Space	Core Bundle	Mapplet	Replaces punctuation symbols including a forward slash, backslash, exclamation point, period, or underscore character with a character space. Also replaces instances of two, three, or four consecutive character spaces with a single space.
c_Merge_Parse_Text	Core Bundle	Cleanse	Merges the Before, After, Overflow and Unparsed columns from the parser. Removes leading and trailing spaces.
c_Remove_Space	Core Bundle	Cleanse	Removes all occurrences of a character space from the input field.
c_Remove_Punctuation	Core Bundle	Cleanse	Removes all occurrences of punctuation from the input field.
c_Uppercase	Core Bundle	Cleanse	Converts the input text to uppercase.

Name	Country	Asset Type	Description
c_remove_ltd_punct_with_space	Core Bundle	Cleanse	Removes limited punctuation and replaces it with a space.
c_merge_remaining_text	Core Bundle	Cleanse	Merges multiple input fields into a single field.
c_Remove_Hyphen	Core Bundle	Cleanse	Removes all hyphens from an input string.
c_Remove_Leading_Zero	Core Bundle	Cleanse	Removes the first character from a number if the character is a zero.
c_Remove_Non_Digit_Chars	Core Bundle	Cleanse	Removes `@` and `+` symbols.
c_Remove_Period_Parentheses	Core Bundle	Cleanse	Removes all occurrences of a period and left and right parentheses in a string.
c_Replace_Limited_Punctuation_w_Space	Core Bundle	Cleanse	Replaces any forward slash, backslash, exclamation point, period, or underscore character with a character space. Also replaces instances of multiple character spaces with a single space.
lbl_Non_Digit_Chars	Core Bundle	Labeler	Labels all non-digit characters consistently.

CHAPTER 9

Accelerating data projects for automation

The Cloud Data Quality Automation Bundle includes data quality rules that support the intelligent automation of rule assignment in cloud profiling. The data quality rules perform cleansing and standardization operations that solve common data quality problems.

Use the bundle to get started with rule specifications and dictionaries that help you to quickly achieve data quality goals. Data Quality can suggest rules for various entities that the accelerator identifies during profiling activity.

Note: Import the **Automation** bundle to a project folder named **CDM**. The folder name enables the suggestion of rules for the source attributes in source types when you create a data profiling task.

Cloud Data Quality automation bundle

Use the assets in the automation bundle to perform common data quality tasks and avoid the overhead of a large team to maintain the Data Quality programs. After the Data Quality process completes, you can see the results in a file or a table.
The following table describes the automation bundle assets:

Name	Type	Description
address_street_line_infa	Dictionary	Contains lists that help interpret the street line of addresses.
calendar_month_frm_number_infa	Dictionary	Contains the names of the calendar months and the numerical equivalent for each month. For example, January and 1.
calendar_months_infa	Dictionary	Contains the names of the calendar months.
company_sufx_abrv_infa	Dictionary	Contains company suffix abbreviations.
company_sufx_long_infa	Dictionary	Contains company suffixes.
countries_infa	Dictionary	Contains the names of all the countries globally.
Email_Domain_Name_Infa	Dictionary	Contains domain names for email addresses. For example, .com , .net, and so on.

Name	Type	Description
numbers_special_characters_infa	Dictionary	Contains a list of numbers and special characters including two spaces.
phone_chars_infa	Dictionary	Contains a list of special characters that can be used as part of a phone number.
timezones_infa	Dictionary	Contains a list of all timezones.
US_Can_States_Infa	Dictionary	Contains the name of the states or provinces of the US and Canada.
us_county_infa	Dictionary	Contains a list of the United States counties.
usa_company_names_std_infa	Dictionary	Contains a list of company name standards in the United States.
usa_firstnames_infa	Dictionary	Contains a list of the United States first names.
usa_nicknames_infa	Dictionary	Contains a list of the United States nicknames.
usa_surnames_infa	Dictionary	Contains a list of the United States surnames.
Validate_Big_Integer	Dictionary	Validates if the input is a big integer.
Validate_Calendar_Month	Rule Specification	Validates if the input is a month.
Validate_Calendar_Quarter	Rule Specification	Validates if the input is a month, and returns the quarter.
Validate_Calendar_Year	Rule Specification	Validates a date, and returns a calendar.
Validate_Character	Rule Specification	Validates if the input is an alphabetical or a special character.
Validate_CompanyName	Rule Specification	Validates if the input is a company name.
Validate_Country	Rule Specification	Validates if the input country is valid.
Validate_Currency	Rule Specification	Validates the input as currency Monetary amount for the price list. - isNullable:true - maximumValue:100000000000000 - minimumValue:0 - is.dataFormat.numeric.shaped
Validate_Date_Creation	Rule Specification	Validates if the input is a date and returns the date in the DD-MM-YYYY format.

Name	Type	Description
Validate_Date_Measurement	Rule Specification	Compares the input date to the current time stamp. Also displays the date and time when the record was migrated. isNullable : true Traits: is.dataFormat.date means.measurement.date is.dataFormat.time means.measurement.time means.measurement.date.creation
Validate_EmailAddress	Rule Specification	Validates if the input is an email address.
Validate_FirstName	Rule Specification	Validates if the input is a first name.
Validate_Integer	Rule Specification	Validates if the input is an integer. - maximumValue: 2147483647 - minimumValue: 2147483648 - Is_nullable: 'True'
Validate_LastName	Rule Specification	Validates if the input is a last name.
Validate_Latitude	Rule Specification	Validates if the input is a valid latitude.
Validate_Longitude	Rule Specification	Validates if the input is a valid longitude.
Validate_MiddleName	Rule Specification	Validates if the input is a middle name.
Validate_Month_frm_Quarter	Rule Specification	Validates the month from a date and returns the name of the month.
Validate_NickName	Rule Specification	Validates if the input is a nickname.
Validate_Phone_Fax	Rule Specification	Validates if the given input is a telephone or a fax number.
Validate_Small_Integer	Rule Specification	Validates if the input string is a small integer.
Validate_Street_Line	Rule Specification	Validates the street line of an address.
Validate_Text	Rule Specification	Validates if the input is an alphabetical text.
Validate_TickerSymbol	Rule Specification	Validates if the input is a valid ticker symbol.
Validate_Timezone	Rule Specification	Validates a given timezone.
Validate_URL	Rule Specification	Validates if the given input is a URL.

Name	Type	Description
Validate_US_Canada_StateName	Rule Specification	Validates if the input is United States state or Canada state or province name.
Validate_USCounty	Rule Specification	Validates if the input is a United States county.

INDEX

B

bundles
  referencing assets 8
  using 6

C

Cloud Application Integration community
  URL 4
Cloud Developer community
  URL 4

D

Data Integration community
  URL 4

I

Informatica Global Customer Support
  contact information 5
Informatica Intelligent Cloud Services
  web site 4

M

maintenance outages 5
mapping tasks
  using bundles 6

R

referencing bundle assets 8

S

status
  Informatica Intelligent Cloud Services 5
system status 5

T

trust site
  description 5

U

upgrade notifications 5

W

web site 4