

EXAMPLE - Double Metaphone Functions

This example illustrates how the following Double Metaphone algorithm functions operate in Trifacta®.

- `DOUBLEMETAPHONE` - Computes a primary and secondary phonetic encoding for an input string. Encodings are returned as a two-element array. See *DOUBLEMETAPHONE Function*.
- `DOUBLEMETAPHONEEQUALS` - Compares two input strings using the Double Metaphone algorithm. Returns `true` if they phonetically match. See *DOUBLEMETAPHONEEQUALS Function*.

Source:

The following table contains some example strings to be compared.

string1	string2	notes
My String	my string	comparison is case-insensitive
judge	jugue	typo
knock	nock	silent letters
white	wite	missing letters
record	record	two different words in English but match the same
pair	pear	these match but are different words.
bookkeeper	book keeper	spaces cause failures in comparison
test1	test123	digits are not compared
the end.	the end....	punctuation differences do not matter.
a elephant	an elephant	a and an are treated differently.

Transformation:

You can use the `DOUBLEMETAPHONE` function to generate phonetic spellings, as in the following:

Transformation Name	New formula
Parameter: Formula type	Single row formula
Parameter: Formula	<code>DOUBLEMETAPHONE(string1)</code>
Parameter: New column name	'dblmeta_s1'

You can compare `string1` and `string2` using the `DOUBLEMETAPHONEEQUALS` function:

Transformation Name	New formula
Parameter: Formula type	Single row formula
Parameter: Formula	<code>DOUBLEMETAPHONEEQUALS(string1, string2, 'normal')</code>
Parameter: New column name	'compare'

Results:

The following table contains some example strings to be compared.

string1	dblmeta_s1	string2	compare	Notes
My String	["MSTRNK", "MSTRNK"]	my string	TRUE	comparison is case-insensitive
judge	["JJ", "AJ"]	judge	TRUE	typo
knock	["NK", "NK"]	nock	TRUE	silent letters
white	["AT", "AT"]	wite	TRUE	missing letters
record	["RKRT", "RKRT"]	record	TRUE	two different words in English but match the same
pair	["PR", "PR"]	pear	TRUE	these match but are different words.
bookkeeper	["PKPR", "PKPR"]	book keeper	FALSE	spaces cause failures in comparison
test1	["TST", "TST"]	test123	TRUE	digits are not compared
the end.	["ONT", "TNT"]	the endâ€¦.	TRUE	punctuation differences do not matter.
a elephant	["ALFNT", "ALFNT"]	an elephant	FALSE	a and an are treated differently.