

Cluster & Edit column "borrower_city_refined"

This feature helps you find groups of different cell values that might be alternative representations of the same thing. For example, "New York" and "New York City" are very likely to refer to the same concept and just have capitalization differences, and "Gödel" and "Godel" probably refer to the same person.

Method key collision

Keying Function ngram-fingerprint

Ngram Size 2

Cluster Size	Row Count	Values in Cluster	Merge?	New Cell Value
3	1955	<ul style="list-style-type: none">Goodyear (1953 rows)Good YearGoodyear	<input type="checkbox"/>	<input type="text" value="Goodyear"/>
2	423	<ul style="list-style-type: none">El Mirage (420 rows)Elmirage (3 rows)	<input type="checkbox"/>	<input type="text" value="El Mirage"/>
2	120	<ul style="list-style-type: none">Youngtown (119 rows)Young Town	<input type="checkbox"/>	<input type="text" value="Youngtown"/>
2	4	<ul style="list-style-type: none">Mc Neal (3 rows)Mcneal	<input type="checkbox"/>	<input type="text" value="Mc Neal"/>
2	43534	<ul style="list-style-type: none">Phoenix (43533 rows)Phoeni X	<input type="checkbox"/>	<input type="text" value="Phoenix"/>
2	812	<ul style="list-style-type: none">Sierra Vista (811 rows)Sierra Vista	<input type="checkbox"/>	<input type="text" value="Sierra Vista"/>