Overview

Dataset Statistics

Number of Variables	3
Number of Rows	1046
Missing Cells	0
Missing Cells (%)	0.0%
Duplicate Rows	0
Duplicate Rows (%)	0.0%
Total Size in Memory	1.0 MB
Average Row Size in Memory	1.0 KB
Variable Types	Categorical: 2 Numerical: 1

Dataset Insights

`seqs` has a high cardinality: 1046 distinct values	High Cardinality
`labels` has constant length 1	Constant Length
`seqs` has all distinct values	Unique

Variables

seqs categorical Show Details	**Approximate Distinct Count**	1046	
	Approximate Unique (%)	**100.0%**	
	Missing	0	
	Missing (%)	0.0%	
	Memory Size	1.0 MB	

seqs

Top 10 of 1046 seqs

Approximate	672	Distinct

	Count		Mean	486.2792
	Approximate Unique (%)	64.2%	Minimum	40
	Missing	0	Maximum	1534
seqs_length	Missing (%)	0.0%	Zeros	0
numerical	Infinite	0	Zeros (%)	0.0%
Show Details	Infinite (%)	0.0%	Negatives	0
	Memory Size	16.3 KB	Negatives (%)	0.0%

	Approximate Distinct Count	2
	Approximate Unique (%)	0.2%
labels	**Missing**	0
categorical	**Missing (%)**	0.0%
Show Details	**Memory Size**	67.4 KB

Interactions

Correlations

Pearson	Spearman	KendallTau

Missing Values

Bar Chart	Spectrum	Heat Map	Dendogram

Report generated with DataPrep