Overview

Dataset Statistics

Number of Variables	3
Number of Rows	539
Missing Cells	0
Missing Cells (%)	0.0%
Duplicate Rows	0
Duplicate Rows (%)	0.0%
Total Size in Memory	293.1 KB
Average Row Size in Memory	556.8 B
Variable Types	Categorical: 2 Numerical: 1

Dataset Insights

`seqs` has a high cardinality: 539 distinct values	High Cardinality
`labels` has constant length 1	Constant Length
`seqs` has all distinct values	Unique

Variables

seqs categorical	**Approximate Distinct Count**	539	
	Approximate Unique (%)	100.0%	
Show Details	**Missing**	0	
	Missing (%)	0.0%	
	Memory Size	288.7 KB	
	Approximate	308	**Distinct**

Top 10 of 539 seqs

	Count		Mean	243.7495
	Approximate Unique (%)	57.1%	Minimum	40
seqs_length	Missing	0	Maximum	432
numerical	Missing (%)	0.0%	Zeros	0
Show Details	Infinite	0	Zeros (%)	0.0%
	Infinite (%)	0.0%	Negatives	0
	Memory Size	8.4 KB	Negatives (%)	0.0%

	Approximate Distinct Count	2
	Approximate Unique (%)	0.4%
labels	Missing	0
categorical	Missing (%)	0.0%
Show Details	Memory Size	34.7 KB

Interactions

Correlations

Pearson	Spearman	KendallTau

Missing Values

Bar Chart	Spectrum	Heat Map	Dendogram

Report generated with DataPrep