

# Overview

Dataset Statistics

Number of Variables	3
Number of Rows	6622
Missing Cells	0
Missing Cells (%)	0.0%
Duplicate Rows	0
Duplicate Rows (%)	0.0%
Total Size in Memory	7.4 MB
Average Row Size in Memory	1.2 KB
Variable Types	Categorical: 2 Numerical: 1

Dataset Insights

seqs\_length

 is skewed

Skewed

seqs

 has a high cardinality: 6622 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	6622
Approximate Unique (%)	100.0%
Missing	0
Missing (%)	0.0%
Memory Size	7.4 MB

seqs

Top 10 of 6622 seqs

Approximate

1491

Distinct

<div>seqs_length</div> <div>numerical</div> <div>Show Details</div>	Count		Mean	557.0979
	Approximate Unique (%)	22.5%	Minimum	40
	Missing	0	Maximum	13100
	Missing (%)	0.0%	Zeros	0
	Infinite	0	Zeros (%)	0.0%
	Infinite (%)	0.0%	Negatives	0
	Memory Size	103.5 KB	Negatives (%)	0.0%
<div>labels</div> <div>categorical</div> <div>Show Details</div>	Approximate Distinct Count			2
	Approximate Unique (%)			0.0%
	Missing			0
	Missing (%)			0.0%
	Memory Size			426.8 KB

# Interactions

# Correlations

# Missing Values

