

DataPrep Report

Overview

Variables =

Interactions

Correlations

Missing Values

Overview

Dataset Statistics

|                            |                                |
|----------------------------|--------------------------------|
| Number of Variables        | 8                              |
| Number of Rows             | 6300                           |
| Missing Cells              | 0                              |
| Missing Cells (%)          | 0.0%                           |
| Duplicate Rows             | 0                              |
| Duplicate Rows (%)         | 0.0%                           |
| Total Size in Memory       | 1.3 MB                         |
| Average Row Size in Memory | 210.2 B                        |
| Variable Types             | Categorical: 3<br>Numerical: 5 |

Dataset Insights

duration\_ms is skewed

id has a high cardinality: 6187 distinct values

name has a high cardinality: 4518 distinct values

id has constant length 22

explicit has constant length 1

popularity has 3331 (52.87%) negatives

duration\_ms has 3597 (57.1%) negatives

Skewed

High Cardinality

High Cardinality

Constant Length

Constant Length

Negatives

Negatives

Variables

Sort by

Feature order

Reverse order

id

categorical

Show Details

|                            |        |
|----------------------------|--------|
| Approximate Distinct Count | 6187   |
| Approximate Unique (%)     | 98.2%  |
| Missing                    | 0      |
| Missing (%)                | 0.0%   |
| Memory Size                | 548100 |

name

categorical

Show Details

|                            |        |
|----------------------------|--------|
| Approximate Distinct Count | 4518   |
| Approximate Unique (%)     | 71.7%  |
| Missing                    | 0      |
| Missing (%)                | 0.0%   |
| Memory Size                | 524056 |

genre

numerical

Show Details

|                            |        |               |      |
|----------------------------|--------|---------------|------|
| Approximate Distinct Count | 126    | Mean          | 62.5 |
| Approximate Unique (%)     | 2.0%   | Minimum       | 0    |
| Missing                    | 0      | Maximum       | 125  |
| Missing (%)                | 0.0%   | Zeros         | 0.8% |
| Infinite                   | 0      | Zeros (%)     | 0.8% |
| Infinite (%)               | 0.0%   | Negatives     | 0    |
| Memory Size                | 100800 | Negatives (%) | 0.0% |

artists

numerical

Show Details

|                            |        |               |           |
|----------------------------|--------|---------------|-----------|
| Approximate Distinct Count | 4651   | Mean          | 2275.0746 |
| Approximate Unique (%)     | 73.8%  | Minimum       | 0         |
| Missing                    | 0      | Maximum       | 4650      |
| Missing (%)                | 0.0%   | Zeros         | 4         |
| Infinite                   | 0      | Zeros (%)     | 0.1%      |
| Infinite (%)               | 0.0%   | Negatives     | 0         |
| Memory Size                | 100800 | Negatives (%) | 0.0%      |

album

numerical

Show Details

|                            |        |               |           |
|----------------------------|--------|---------------|-----------|
| Approximate Distinct Count | 4949   | Mean          | 2449.1551 |
| Approximate Unique (%)     | 78.6%  | Minimum       | 0         |
| Missing                    | 0      | Maximum       | 4948      |
| Missing (%)                | 0.0%   | Zeros         | 1         |
| Infinite                   | 0      | Zeros (%)     | 0.0%      |
| Infinite (%)               | 0.0%   | Negatives     | 0         |
| Memory Size                | 100800 | Negatives (%) | 0.0%      |

popularity

numerical

Show Details

|                            |        |               |                           |
|----------------------------|--------|---------------|---------------------------|
| Approximate Distinct Count | 90     | Mean          | -9.0228×10 <sup>-17</sup> |
| Approximate Unique (%)     | 1.4%   | Minimum       | -1.5418                   |
| Missing                    | 0      | Maximum       | 2.9701                    |
| Missing (%)                | 0.0%   | Zeros         | 0                         |
| Infinite                   | 0      | Zeros (%)     | 0.0%                      |
| Infinite (%)               | 0.0%   | Negatives     | 3331                      |
| Memory Size                | 100800 | Negatives (%) | 52.9%                     |

duration\_ms

numerical

Show Details

|                            |        |               |                           |
|----------------------------|--------|---------------|---------------------------|
| Approximate Distinct Count | 5566   | Mean          | -1.0827×10 <sup>-16</sup> |
| Approximate Unique (%)     | 88.3%  | Minimum       | -1.4278                   |
| Missing                    | 0      | Maximum       | 28.0846                   |
| Missing (%)                | 0.0%   | Zeros         | 0                         |
| Infinite                   | 0      | Zeros (%)     | 0.0%                      |
| Infinite (%)               | 0.0%   | Negatives     | 3597                      |
| Memory Size                | 100800 | Negatives (%) | 57.1%                     |

explicit

categorical

Show Details

|                            |        |
|----------------------------|--------|
| Approximate Distinct Count | 2      |
| Approximate Unique (%)     | 0.0%   |
| Missing                    | 0      |
| Missing (%)                | 0.0%   |
| Memory Size                | 415800 |

id

count

Top 10 of 6187 id

name

count

Top 10 of 4518 name

genre

frequency

genre

artists

frequency

artists

album

frequency

album

popularity

frequency

popularity

duration\_ms

frequency

duration\_ms

explicit

count

explicit

Interactions

X-Axis

genre

Y-Axis

genre

Scatter Plot

genre

genre

Correlations

Pearson

Spearman

KendallTau

genre

artists

album

popularity

duration\_ms

explicit

genre

artists

album

popularity

duration\_ms

explicit

genre

artists

album

popularity

duration\_ms

explicit

Missing Values

Bar Chart

Spectrum

Heat Map

Dendrogram

Present

Missing

id

name

genre

artists

album

popularity

duration\_ms

explicit

id

name

genre

artists

album

popularity

duration\_ms

explicit

Report generated with DataPrep