# Overview

## Dataset statistics

| | |
|---|---|
| **Number of variables** | 2 |
| **Number of observations** | 265 |
| **Missing cells** | 1 |
| **Missing cells (%)** | 0.2% |
| **Duplicate rows** | 0 |
| **Duplicate rows (%)** | 0.0% |
| **Total size in memory** | 20.2 KiB |
| **Average record size in memory** | 77.9 B |

## Variable types

| | |
|---|---|
| **Text** | 1 |
| **Numeric** | 1 |

## Alerts

Country Name has unique values                                    Unique

## Reproduction

| | |
|---|---|
| **Analysis started** | 2023-09-17 21:33:39.925138 |
| **Analysis finished** | 2023-09-17 21:33:40.446930 |
| **Duration** | 0.52 seconds |
| **Software version** | ydata-profiling vv4.5.1 (https://github.com/ydataai/ydata-profiling) |
| **Download configuration** | config.json (data:text/plain;charset=utf-8,%7B%22title%22%3A%20%22Country%20Population%20Report%22%2 |

# Variables

Select Columns ⌄

## Country Name
Text

UNIQUE

| | |
|---|---|
| **Distinct** | 265 |
| **Distinct (%)** | 100.0% |
| **Missing** | 0 |
| **Missing (%)** | 0.0% |
| **Memory size** | 18.1 KiB |

## Length

| | |
|---|---|
| **Max length** | 52 |
| **Median length** | 44 |
| **Mean length** | 12.407547 |
| **Min length** | 4 |

## Characters and Unicode

| | | |
|---|---|---|
| **Total characters** | 3288 | |
| **Distinct characters** | 60 | |
| **Distinct categories** | 7 (https://en.wikipedia.org/wiki/ Unicode_character_property#General_Category) | ? |
| **Distinct scripts** | 2 (https://en.wikipedia.org/wiki/ Script_(Unicode)#List_of_scripts_in_Unicode) | ? |
| **Distinct blocks** | 1 (https://en.wikipedia.org/wiki/Unicode_block) | ? |

The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.

## Unique

| | | |
|---|---|---|
| **Unique** | 265 | ? |
| **Unique (%)** | 100.0% | |

## Sample

| | |
|---|---|
| **1st row** | Africa Eastern and Southern |
| **2nd row** | Afghanistan |
| **3rd row** | Africa Western and Central |
| **4th row** | Angola |
| **5th row** | Albania |

| Value | Count | Frequency (%) |
|---|---|---|
| | 20 | 4.0% |
| and | 12 | 2.4% |
| income | 11 | 2.2% |
| ida | 10 | 2.0% |
| africa | 9 | 1.8% |
| islands | 9 | 1.8% |
| asia | 8 | 1.6% |
| ibrd | 8 | 1.6% |
| countries | 7 | 1.4% |
| middle | 7 | 1.4% |
| Other values (309) | 402 | 79.9% |

## Most occurring characters

| Value | Count | Frequency (%) |
| --- | ---: | ---: |
| a | 387 | 11.8% |
| i | 260 | 7.9% |
| | 238 | 7.2% |
| e | 229 | 7.0% |
| n | 222 | 6.8% |
| r | 175 | 5.3% |
| o | 157 | 4.8% |
| t | 130 | 4.0% |
| s | 120 | 3.6% |
| l | 114 | 3.5% |
| Other values (50) | 1256 | 38.2% |

## Most occurring categories

| Value | Count | Frequency (%) |
|---|---|---|
| Lowercase Letter | 2487 | 75.6% |
| Uppercase Letter | 471 | 14.3% |
| Space Separator | 238 | 7.2% |
| Other Punctuation | 53 | 1.6% |
| Open Punctuation | 15 | 0.5% |
| Close Punctuation | 15 | 0.5% |
| Dash Punctuation | 9 | 0.3% |

## Most frequent character per category

*Lowercase Letter*

| Value | Count | Frequency (%) |
|---|---|---|
| a | 387 | 15.6% |
| i | 260 | 10.5% |
| e | 229 | 9.2% |
| n | 222 | 8.9% |
| r | 175 | 7.0% |
| o | 157 | 6.3% |
| t | 130 | 5.2% |
| s | 120 | 4.8% |
| l | 114 | 4.6% |
| d | 113 | 4.5% |
| Other values (16) | 580 | 23.3% |

*Uppercase Letter*

| Value | Count | Frequency (%) |
|---|---|---|
| A | 51 | 10.8% |
| S | 45 | 9.6% |
| I | 39 | 8.3% |
| C | 36 | 7.6% |
| B | 32 | 6.8% |
| M | 29 | 6.2% |
| R | 29 | 6.2% |
| D | 28 | 5.9% |
| E | 24 | 5.1% |
| L | 21 | 4.5% |
| Other values (15) | 137 | 29.1% |

*Other Punctuation*

| Value | Count | Frequency (%) |
|---|---|---|
| & | 20 | 37.7% |
| . | 17 | 32.1% |
| , | 13 | 24.5% |
| ' | 2 | 3.8% |
| : | 1 | 1.9% |

*Space Separator*

| Value | Count | Frequency (%) |
|---|---|---|
| | 238 | 100.0% |

*Open Punctuation*

| Value | Count | Frequency (% |
|---|---|---|
| ( | 15 | 100.0% |

*Close Punctuation*

| Value | Count | Frequency (% |
|---|---|---|
| ) | 15 | 100.0% |

*Dash Punctuation*

| Value | Count | Frequency (% |
|---|---|---|
| - | 9 | 100.0% |

# Most occurring scripts

| Value | Count | Frequency (%) |
|---|---|---|
| Latin | 2958 | 90.0% |
| Common | 330 | 10.0% |

# Most frequent character per script

*Latin*

| Value | Count | Frequency (%) |
|---|---|---|
| a | 387 | 13.1% |
| i | 260 | 8.8% |
| e | 229 | 7.7% |
| n | 222 | 7.5% |
| r | 175 | 5.9% |
| o | 157 | 5.3% |
| t | 130 | 4.4% |
| s | 120 | 4.1% |
| l | 114 | 3.9% |
| d | 113 | 3.8% |
| Other values (41) | 1051 | 35.5% |

*Common*

| Value | Count | Frequency (%) |
|---|---|---|
| | 238 | 72.1% |
| & | 20 | 6.1% |
| . | 17 | 5.2% |

| Value | Count | Frequency (%) |
|---|---|---|
| ( | 15 | 4.5% |
| ) | 15 | 4.5% |
| , | 13 | 3.9% |
| - | 9 | 2.7% |
| ' | 2 | 0.6% |
| : | 1 | 0.3% |

## Most occurring blocks

| Value | Count | Frequency (%) |
|-------|-------|---------------|
| ASCII | 3288 | 100.0% |

## Most frequent character per block

*ASCII*

| Value | Count | Frequency (%) |
|-------|-------|---------------|
| a | 387 | 11.8% |
| i | 260 | 7.9% |
|  | 238 | 7.2% |
| e | 229 | 7.0% |
| n | 222 | 6.8% |
| r | 175 | 5.3% |
| o | 157 | 4.8% |
| t | 130 | 4.0% |
| s | 120 | 3.6% |
| l | 114 | 3.5% |
| Other values (50) | 1256 | 38.2% |

## 2022

Real number (ℝ)

| | |
|---|---|
| **Distinct** | 262 |
| **Distinct (%)** | 99.2% |
| **Missing** | 1 |
| **Missing (%)** | 0.4% |
| **Infinite** | 0 |
| **Infinite (%)** | 0.0% |
| **Mean** | 0.92286234 |
| **Minimum** | -14.188452 |
| **Maximum** | 3.7129878 |
| **Zeros** | 0 |
| **Zeros (%)** | 0.0% |
| **Negative** | 45 |
| **Negative (%)** | 17.0% |
| **Memory size** | 2.2 KiB |
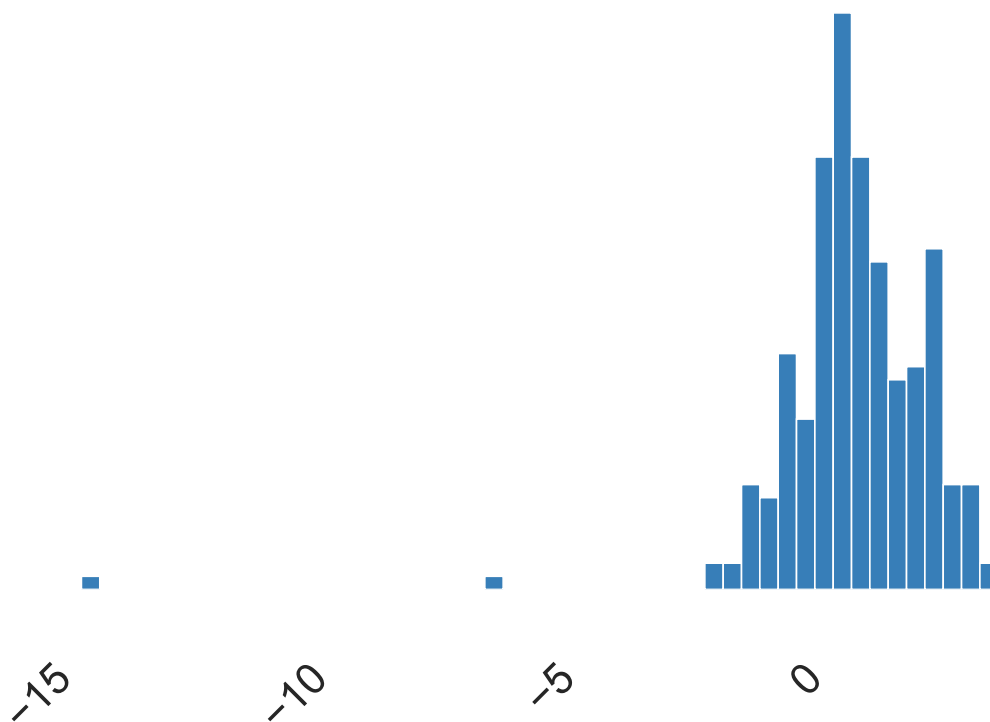
## Quantile statistics

| | |
|---|---|
| **Minimum** | -14.188452 |
| **5-th percentile** | -0.94852237 |
| **Q1** | 0.29076453 |
| **median** | 0.89684588 |
| **Q3** | 1.8157101 |
| **95-th percentile** | 2.701213 |
| **Maximum** | 3.7129878 |
| **Range** | 17.901439 |
| **Interquartile range (IQR)** | 1.5249455 |

## Descriptive statistics

| | |
|---|---|
| **Standard deviation** | 1.5100987 |
| **Coefficient of variation (CV)** | 1.6363206 |
| **Kurtosis** | 38.65979 |

| | |
|---|---|
| **Mean** | 0.92286234 |
| **Median Absolute Deviation (MAD)** | 0.72412936 |
| **Skewness** | -4.1386192 |
| **Sum** | 243.63566 |
| **Variance** | 2.280398 |
| **Monotonicity** | Not monotonic |
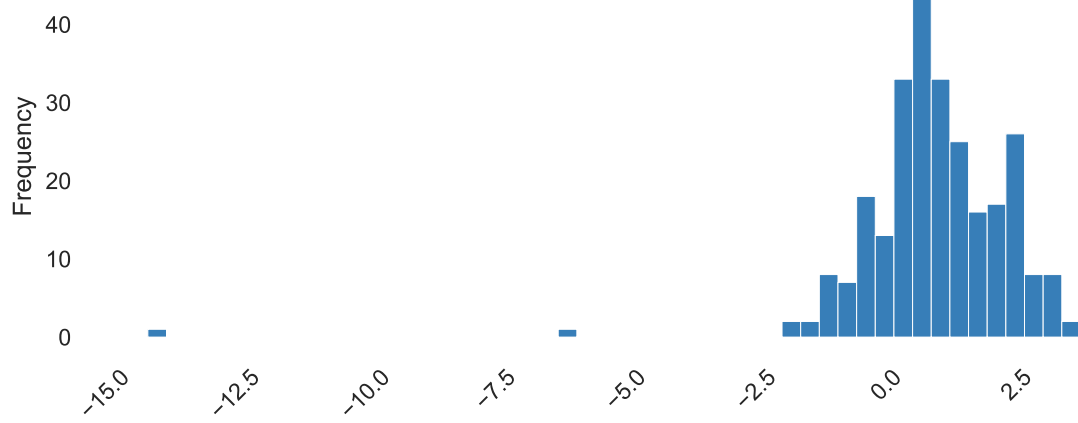
**Histogram with fixed size bins** (bins=50)

| Value | Count | Frequency (% |
|---|---|---|
| 2.540476689 | 2 | 0.8% |
| 0.916782671 | 2 | 0.8% |
| 2.540937372 | 1 | 0.4% |
| 2.380007155 | 1 | 0.4% |
| 1.247087134 | 1 | 0.4% |
| 1.692246601 | 1 | 0.4% |
| 0.898395221 | 1 | 0.4% |
| 0.965187158 | 1 | 0.4% |
| 1.418278462 | 1 | 0.4% |
| 3.712987818 | 1 | 0.4% |
| Other values (252) | 252 | 95.1% |

| Value | Count | Frequency (%) |
| --- | --- | --- |
| -14.18845165 | 1 | 0.4% |
| -6.187252982 | 1 | 0.4% |
| -1.856912493 | 1 | 0.4% |
| -1.706495481 | 1 | 0.4% |
| -1.569658341 | 1 | 0.4% |
| -1.35587444 | 1 | 0.4% |
| -1.289800065 | 1 | 0.4% |
| -1.261613265 | 1 | 0.4% |
| -1.150513985 | 1 | 0.4% |
| -1.150468925 | 1 | 0.4% |

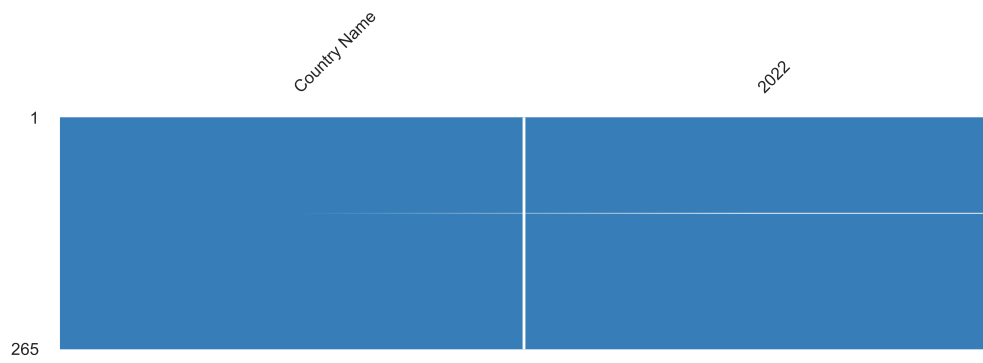| Value | Count | Frequency (% |
|---|---|---|
| 3.712987818 | 1 | 0.4% |
| 3.68690356 | 1 | 0.4% |
| 3.308620524 | 1 | 0.4% |
| 3.215898096 | 1 | 0.4% |
| 3.197835105 | 1 | 0.4% |
| 3.115022186 | 1 | 0.4% |
| 3.096752671 | 1 | 0.4% |
| 3.095209254 | 1 | 0.4% |
| 3.069384352 | 1 | 0.4% |
| 2.99862764 | 1 | 0.4% |

# Interactions
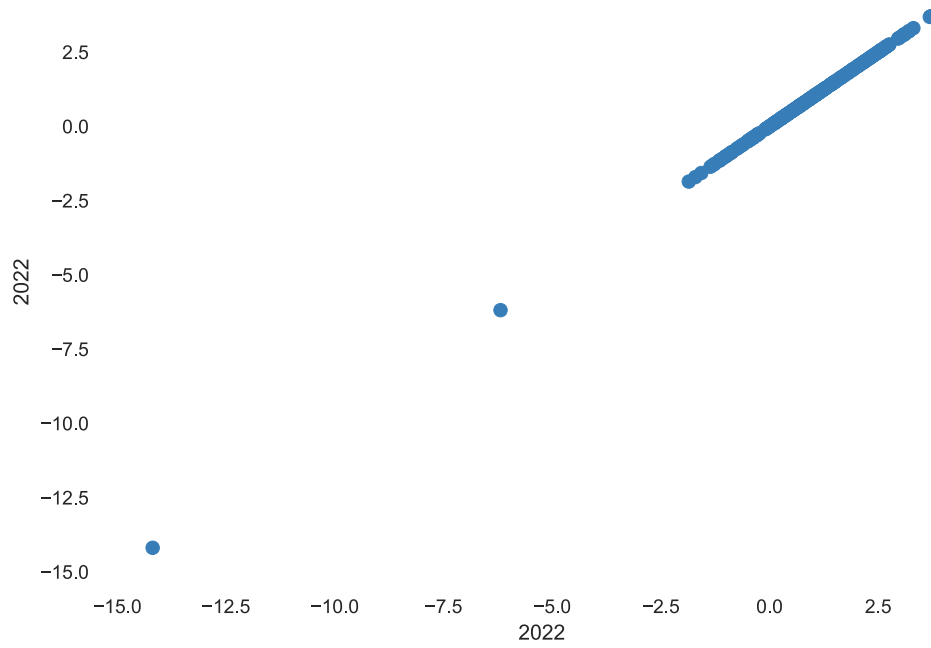
# Missing values
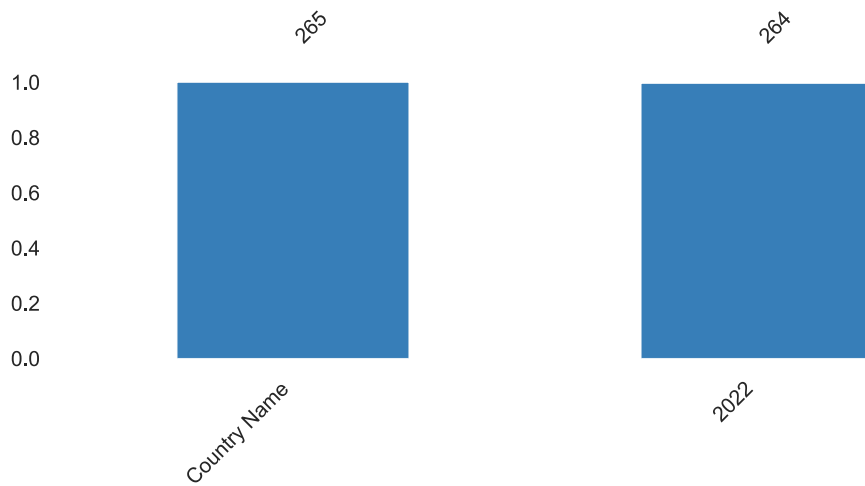


A simple visualization of nullity by column.

Nullity matrix is a data-dense display which lets you quickly visually pick out patterns in data completion.

# Sample

|    | Country Name                 | 2022      |
|----|------------------------------|-----------|
| 1  | Africa Eastern and Southern  | 2.540937  |
| 2  | Afghanistan                  | 2.534498  |
| 3  | Africa Western and Central   | 2.539799  |
| 4  | Angola                       | 3.096753  |
| 5  | Albania                      | -1.289800 |
| 6  | Andorra                      | 0.994607  |
| 7  | Arab World                   | 1.788339  |
| 8  | United Arab Emirates         | 0.808075  |
| 9  | Argentina                    | 0.925835  |
| 10 | Armenia                      | -0.377102 |

| | Country Name | 2022 |
|---|---|---|
| **256** | Virgin Islands (U.S.) | -0.432596 |
| **257** | Vietnam | 0.734794 |
| **258** | Vanuatu | 2.354427 |
| **259** | World | 0.798516 |
| **260** | Samoa | 1.640310 |
| **261** | Kosovo | -1.355874 |
| **262** | Yemen, Rep. | 2.144628 |
| **263** | South Africa | 0.841058 |
| **264** | Zambia | 2.758032 |
| **265** | Zimbabwe | 2.024036 |