dados\_sim

Autogenerated data summary from dataMaid

2023-08-28 00:47:09.424657

# Data report overview

The dataset examined has the following dimensions:

| Feature | Result |
| --- | --- |
| Number of observations | 342213 |
| Number of variables | 17 |

### Checks performed

The following variable checks were performed, depending on the data type of each variable:

|  | character | factor | labelled | haven labelled | numeric | integer | logical | Date |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Identify miscoded missing values | × | × | × | × | × | × |  | × |
| Identify prefixed and suffixed whitespace | × | × | × | × |  |  |  |  |
| Identify levels with < 6 obs. | × | × | × | × |  |  |  |  |
| Identify case issues | × | × | × | × |  |  |  |  |
| Identify misclassified numeric or integer variables | × | × | × | × |  |  |  |  |
| Identify outliers |  |  |  |  | × | × |  | × |

Please note that all numerical values in the following have been rounded to 2 decimals.

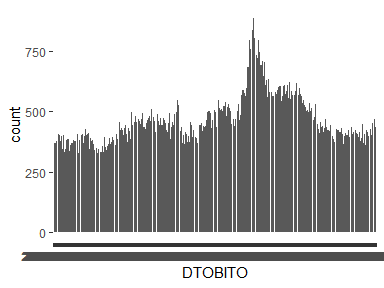
# Summary table

|  | Variable class | # unique values | Missing observations | Any problems? |
| --- | --- | --- | --- | --- |
| [DTOBITO](#dtobito) | character | 731 | 0.00 % |  |
| [DTNASC](#dtnasc) | character | 33709 | 0.14 % | × |
| [SEXO](#sexo) | character | 3 | 0.02 % |  |
| [CODMUNRES](#codmunres) | character | 854 | 0.00 % | × |
| [IDADEminutos](#idademinutos) | character | 60 | 99.85 % | × |
| [IDADEhoras](#idadehoras) | character | 25 | 99.74 % | × |
| [IDADEdias](#idadedias) | character | 31 | 99.29 % | × |
| [IDADEmeses](#idademeses) | character | 12 | 99.61 % | × |
| [IDADEanos](#idadeanos) | character | 121 | 1.55 % | × |
| [munResStatus](#munresstatus) | character | 2 | 0.00 % |  |
| [munResTipo](#munrestipo) | character | 2 | 0.00 % |  |
| [munResNome](#munresnome) | character | 854 | 0.00 % |  |
| [munResUf](#munresuf) | character | 1 | 0.00 % | × |
| [munResLat](#munreslat) | character | 853 | 0.09 % | × |
| [munResLon](#munreslon) | character | 854 | 0.09 % | × |
| [munResAlt](#munresalt) | character | 551 | 0.09 % | × |
| [munResArea](#munresarea) | character | 853 | 0.09 % | × |

# Variable list

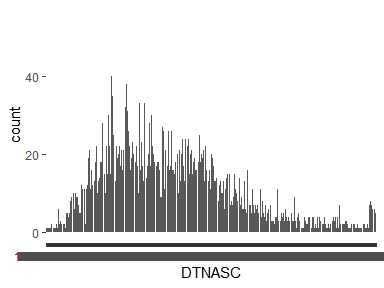
## DTOBITO

| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 731 |
| Mode | “2021-03-29” |



## DTNASC

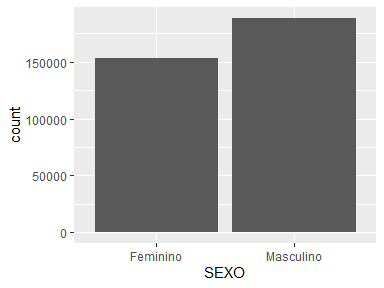
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 469 (0.14 %) |
| Number of unique values | 33708 |
| Mode | “1936-05-20” |



* Note that the following levels have at most five observations: "1899-04-18", "1900-12-20", "1902-01-06", "1902-02-14", "1902-11-09", …, "2021-12-26", "2021-12-28", "2021-12-29", "2021-12-30", "2021-12-31" (13307 values omitted).

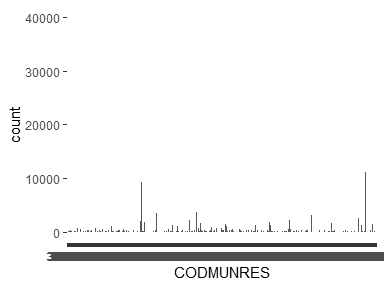
## SEXO

| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 68 (0.02 %) |
| Number of unique values | 2 |
| Mode | “Masculino” |



## CODMUNRES

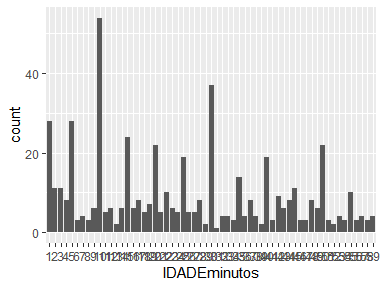
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 854 |
| Mode | “310620” |



* Note: The variable consists exclusively of numbers and takes a lot of different values. Is it perhaps a misclassified numeric variable?

## IDADEminutos

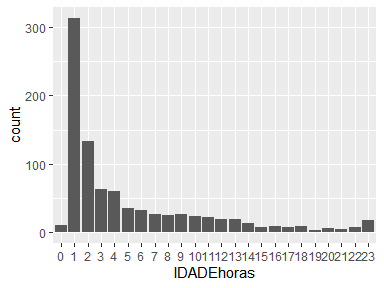
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 341686 (99.85 %) |
| Number of unique values | 59 |
| Mode | “10” |



* Note that the following levels have at most five observations: "11", "13", "18", "21", "24", …, "58", "59", "6", "7", "8" (19 values omitted).
* Note: The variable consists exclusively of numbers and takes a lot of different values. Is it perhaps a misclassified numeric variable?

## IDADEhoras

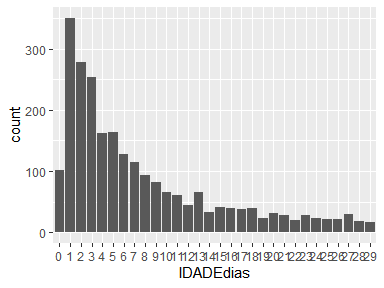
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 341314 (99.74 %) |
| Number of unique values | 24 |
| Mode | “1” |



* Note that the following levels have at most five observations: "19", "21".
* Note: The variable consists exclusively of numbers and takes a lot of different values. Is it perhaps a misclassified numeric variable?

## IDADEdias

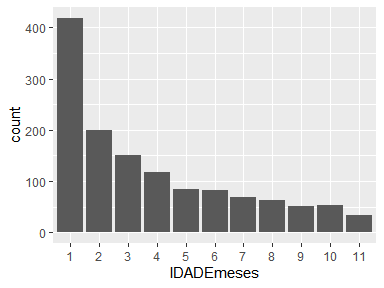
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 339795 (99.29 %) |
| Number of unique values | 30 |
| Mode | “1” |



* Note: The variable consists exclusively of numbers and takes a lot of different values. Is it perhaps a misclassified numeric variable?

## IDADEmeses

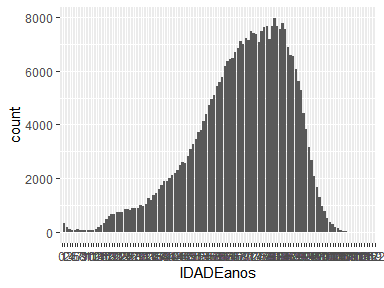
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 340889 (99.61 %) |
| Number of unique values | 11 |
| Mode | “1” |



* The following suspected missing value codes enter as regular values: "8", "9".

## IDADEanos

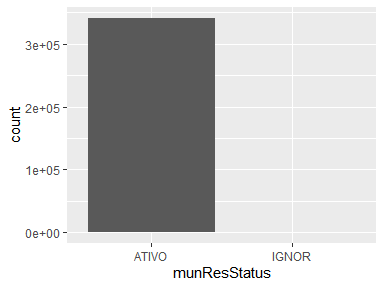
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 5301 (1.55 %) |
| Number of unique values | 120 |
| Mode | “81” |



* Note that the following levels have at most five observations: "0", "115", "116", "117", "119", "122".
* Note: The variable consists exclusively of numbers and takes a lot of different values. Is it perhaps a misclassified numeric variable?

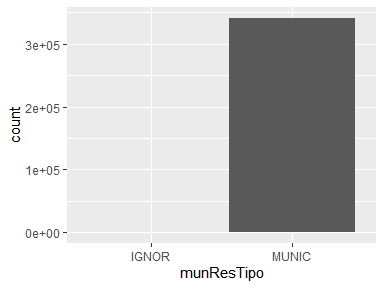
## munResStatus

| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 2 |
| Mode | “ATIVO” |



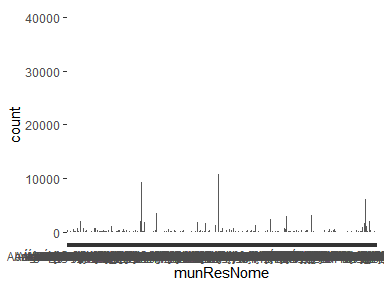
## munResTipo

| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 2 |
| Mode | “MUNIC” |



## munResNome

| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 854 |
| Mode | “Belo Horizonte” |

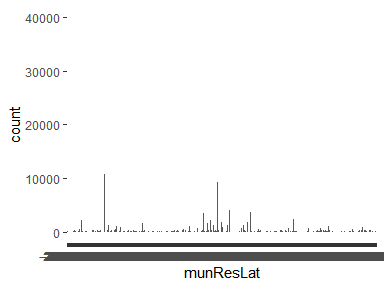


## munResUf

* The variable only takes one (non-missing) value: "Minas Gerais". The variable contains 0 % missing observations.

## munResLat

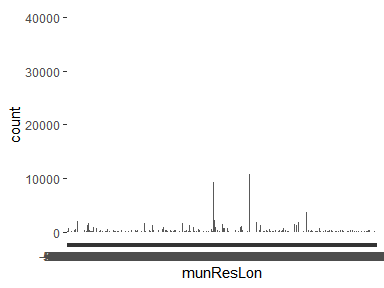
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 313 (0.09 %) |
| Number of unique values | 852 |
| Mode | “-19.93752” |



* Note: The variable consists exclusively of numbers and takes a lot of different values. Is it perhaps a misclassified numeric variable?

## munResLon

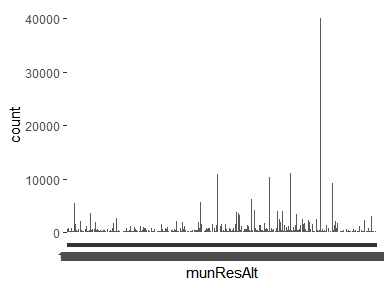
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 313 (0.09 %) |
| Number of unique values | 853 |
| Mode | “-43.92645” |



* Note: The variable consists exclusively of numbers and takes a lot of different values. Is it perhaps a misclassified numeric variable?

## munResAlt

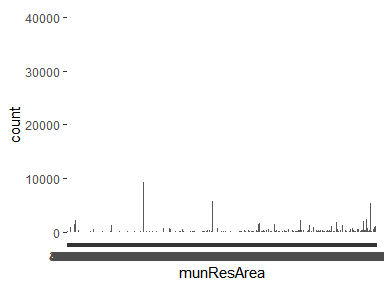
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 313 (0.09 %) |
| Number of unique values | 550 |
| Mode | “938” |



* Note: The variable consists exclusively of numbers and takes a lot of different values. Is it perhaps a misclassified numeric variable?

## munResArea

| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 313 (0.09 %) |
| Number of unique values | 852 |
| Mode | “331.401” |



* Note: The variable consists exclusively of numbers and takes a lot of different values. Is it perhaps a misclassified numeric variable?

Report generation information:

* Created by: Could not determine from system (username: Unknown)
* Report creation time: seg ago 28 2023 00:47:46
* Report was run from directory: C:/Users/caios/Documents/GitHub/mqagp
* dataMaid v1.4.1 [Pkg: 2021-10-08 from CRAN (R 4.3.1)]
* R version 4.3.1 (2023-06-16 ucrt).
* Platform: x86\_64-w64-mingw32/x64 (64-bit)(America/Sao\_Paulo).
* Function call: makeDataReport(data = dados\_sim, replace = TRUE)