

Overview

Brought to you by [YData](#)

Overview

Alerts 17

Reproduction

Alerts

Dataset has 12 (< 0.1%) duplicate rows	Duplicates
cons.conf.idx is highly overall correlated with month	High correlation
cons.price.idx is highly overall correlated with contact and 2 other fields	High correlation
contact is highly overall correlated with cons.price.idx and 2 other fields	High correlation
emp.var.rate is highly overall correlated with cons.price.idx and 3 other fields	High correlation
euribor3m is highly overall correlated with emp.var.rate and 2 other fields	High correlation
housing is highly overall correlated with loan	High correlation
loan is highly overall correlated with housing	High correlation
month is highly overall correlated with cons.conf.idx and 5 other fields	High correlation
nr.employed is highly overall correlated with contact and 3 other fields	High correlation
pdays is highly overall correlated with poutcome and 1 other fields	High correlation
poutcome is highly overall correlated with pdays and 1 other fields	High correlation
previous is highly overall correlated with pdays and 1 other fields	High correlation
default is highly imbalanced (53.3%)	Imbalance
loan is highly imbalanced (51.3%)	Imbalance
poutcome is highly imbalanced (56.8%)	Imbalance
previous has 35563 (86.3%) zeros	Zeros

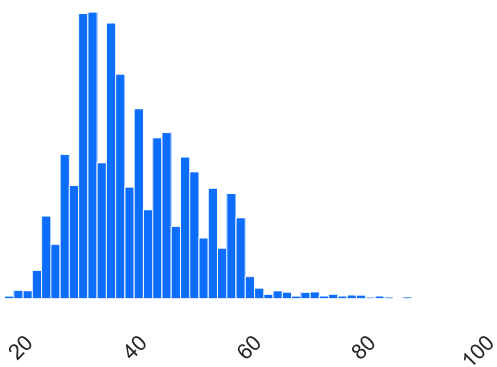
Variables

Select Columns ▾

age

Real number (\mathbb{R})

Distinct	78	Minimum	17
Distinct (%)	0.2%	Maximum	98
Missing	0	Zeros	0
Missing (%)	0.0%	Zeros (%)	0.0%
Infinite	0	Negative	0
Infinite (%)	0.0%	Negative (%)	0.0%
Mean	40.02406	Memory size	321.9 KiB



More details

Statistics

Histogram

Common values

Extreme values

Quantile statistics

Minimum	17
5-th percentile	26
Q1	32
median	38
Q3	47
95-th percentile	58
Maximum	98
Range	81
Interquartile range (IQR)	15

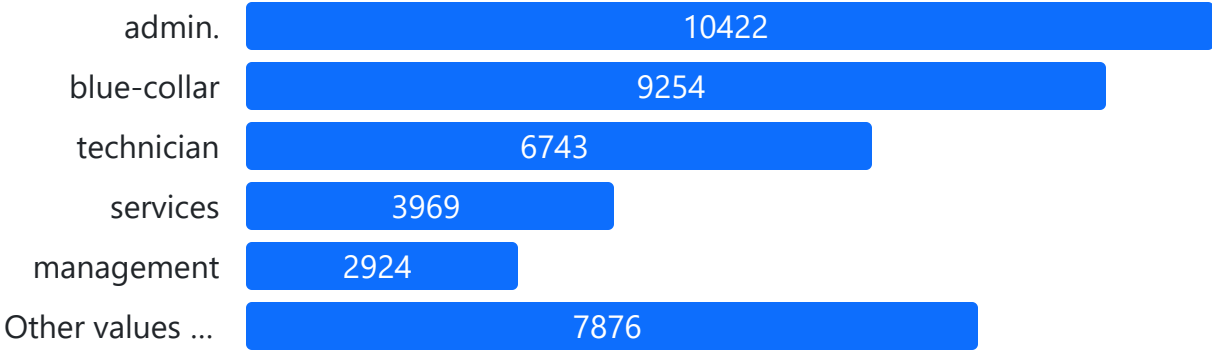
Descriptive statistics

Standard deviation	10.42125
Coefficient of variation (CV)	0.26037463
Kurtosis	0.79131153
Mean	40.02406
Median Absolute Deviation (MAD)	7
Skewness	0.78469682
Sum	1648511
Variance	108.60245
Monotonicity	Not monotonic

job

Categorical

Distinct	12
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	321.9 KiB



More details

Overview

Categories

Words

Characters

Length

Max length	13
Median length	12
Mean length	8.9552297
Min length	6

Characters and Unicode

Total characters	368848
Distinct characters	24
Distinct categories	1
Distinct scripts	1
Distinct blocks	1

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

Unique

Unique	0
Unique (%)	0.0%

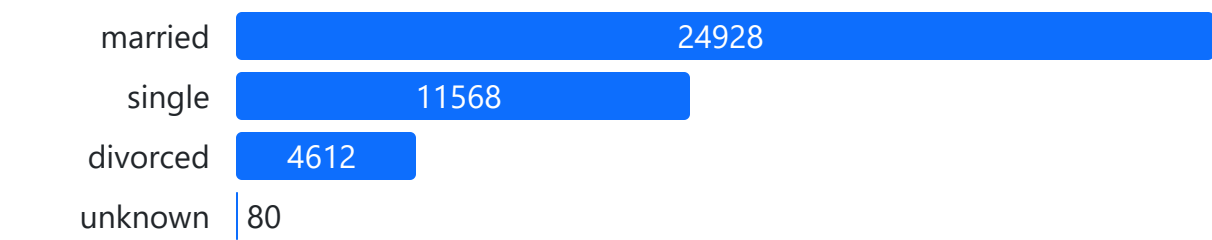
Sample

1st row	housemaid
2nd row	services
3rd row	services
4th row	admin.
5th row	services

marital

Categorical

Distinct	4
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	321.9 KiB



More details

OverviewCategoriesWordsCharacters

Length

Max length	8
Median length	7
Mean length	6.8311159
Min length	6

Characters and Unicode

Total characters	281360
Distinct characters	16
Distinct categories	1?
Distinct scripts	1?
Distinct blocks	1?

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

Unique

Unique	0?
Unique (%)	0.0%

Sample

1st row	married
2nd row	married
3rd row	married
4th row	married
5th row	married

education

Categorical

Distinct

8

Distinct (%)

< 0.1%

Missing

0

Missing (%)

0.0%

Memory size

321.9 KiB

university.de...

12168

high.school

9515

basic.9y

6045

professional...

5243

basic.4y

4176

Other values ...

4041

More details

OverviewCategoriesWordsCharacters

Length

Max length	19
Median length	17
Mean length	12.71096
Min length	7

Characters and Unicode

Total characters	523539
Distinct characters	25
Distinct categories	1?
Distinct scripts	1?

Unique

Unique	0?
Unique (%)	0.0%

Sample

1st row	basic.4y
2nd row	high.school
3rd row	high.school
4th row	basic.6y

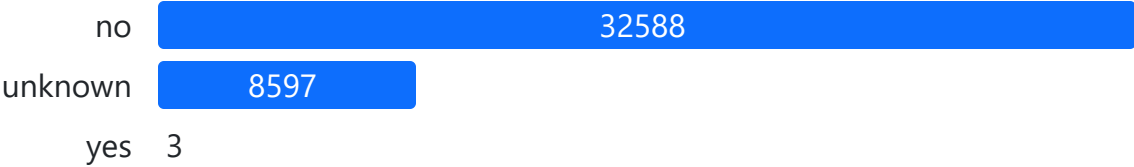
	<div>Distinct blocks1?</div>		<div>5th rowhigh.school</div>
<div>The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.</div>			

default

Categorical

Imbalance

Distinct	3
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	321.9 KiB



More details

Overview

Categories

Words

Characters

Length

Max length	7
Median length	2
Mean length	3.043702
Min length	2

Characters and Unicode

Total characters	125364
Distinct characters	8
Distinct categories	1?
Distinct scripts	1?
Distinct blocks	1?

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

Unique

Unique	0?
Unique (%)	0.0%

Sample

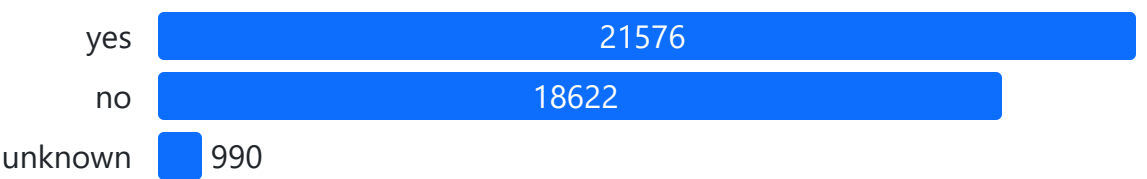
1st row	no
2nd row	unknown
3rd row	no
4th row	no
5th row	no

housing

Categorical

High correlation

Distinct	3
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	321.9 KiB



More details

Overview

Categories

Words

Characters

Length

Max length	7
Median length	3
Mean length	2.6440225
Min length	2

Characters and Unicode

Total characters	108902
Distinct characters	8
Distinct categories	1
Distinct scripts	1
Distinct blocks	1

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

Unique

Unique	0
Unique (%)	0.0%

Sample

1st row	no
2nd row	no
3rd row	yes
4th row	no
5th row	no

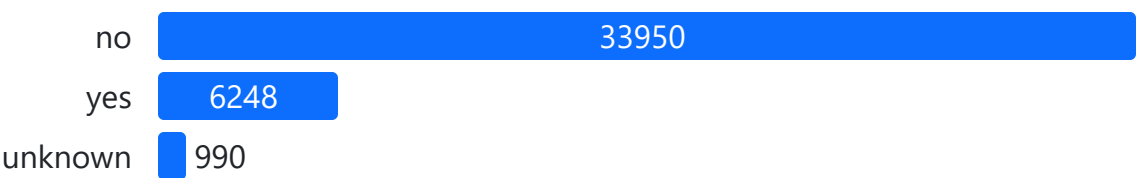
loan

Categorical

High correlation

Imbalance

Distinct	3
Distinct (%)	< 0.1%



Missing

0

Missing (%)

0.0%

Memory size

321.9 KiB

More details

Overview

Categories

Words

Characters

Length

Max length

7

Median length

2

Mean length

2.2718753

Min length

2

Characters and Unicode

Total characters

93574

Distinct characters

8

Distinct categories

1

?

Distinct scripts

1

?

Distinct blocks

1

?

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

Unique

Unique

0

?

Unique (%)

0.0%

Sample

1st row

no

2nd row

no

3rd row

no

4th row

no

5th row

yes

<div>contact</div> <div>Categorical</div> <div>High correlation</div>	
Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	321.9 KiB
<div>cellular26144</div> <div>telephone15044</div>	
<div>More details</div>	
<div>OverviewCategoriesWordsCharacters</div>	

Length		Characters and Unicode		Unique		Sample	
Max length	9	Total characters	344548	Unique	0	1st row	telephone
Median length	8	Distinct characters	11	Unique (%)	0.0%	2nd row	telephone
Mean length	8.365252	Distinct categories	1			3rd row	telephone
Min length	8	Distinct scripts	1			4th row	telephone
		Distinct blocks	1			5th row	telephone
The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.							

month

Categorical

High correlation

Distinct	10
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	321.9 KiB

may

13769

jul

7174

aug

6178

jun

5318

nov

4101

Other values ...

4648

More details

Overview

Categories

Words

Characters

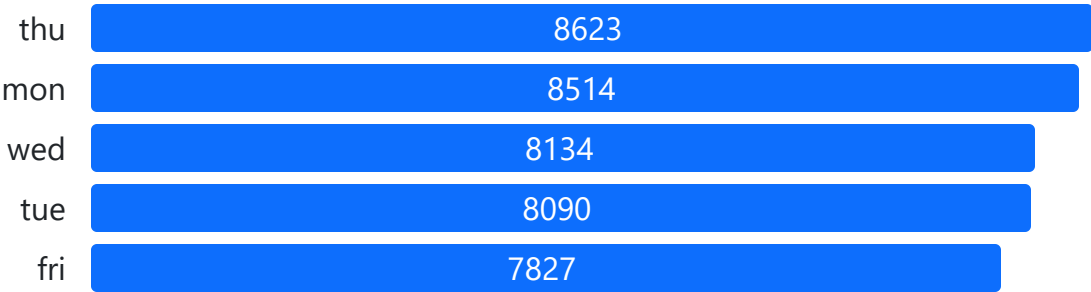
Length		Characters and Unicode		Unique		Sample	
Max length	3	Total characters	123564	Unique	0	1st row	may
Median length	3	Distinct characters	17	Unique (%)	0.0%	2nd row	may
Mean length	3	Distinct categories	1			3rd row	may
Min length	3	Distinct scripts	1			4th row	may

	Distinct blocks	1	<div>?</div>		5th row	may
The Unicode Standard assigns character properties to each code point, which can be used to analyse textual variables.						

day_of_week

Categorical

Distinct	5
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	321.9 KiB



More details

Overview

Categories

Words

Characters

Length

Max length	3
Median length	3
Mean length	3
Min length	3

Characters and Unicode

Total characters	123564	
Distinct characters	12	
Distinct categories	1	<div>?</div>
Distinct scripts	1	<div>?</div>
Distinct blocks	1	<div>?</div>

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

Unique

Unique	0	<div>?</div>
Unique (%)	0.0%	

Sample

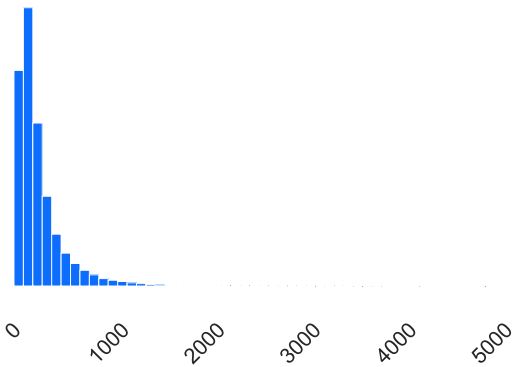
1st row	mon
2nd row	mon
3rd row	mon
4th row	mon
5th row	mon

duration

Real number (\mathbb{R})

Distinct	1544
Distinct (%)	3.7%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	258.28501

Minimum	0
Maximum	4918
Zeros	4
Zeros (%)	< 0.1%
Negative	0
Negative (%)	0.0%
Memory size	321.9 KiB



More details

Statistics

Histogram

Common values

Extreme values

Quantile statistics

Minimum	0
5-th percentile	36
Q1	102
median	180
Q3	319
95-th percentile	752.65
Maximum	4918
Range	4918
Interquartile range (IQR)	217

Descriptive statistics

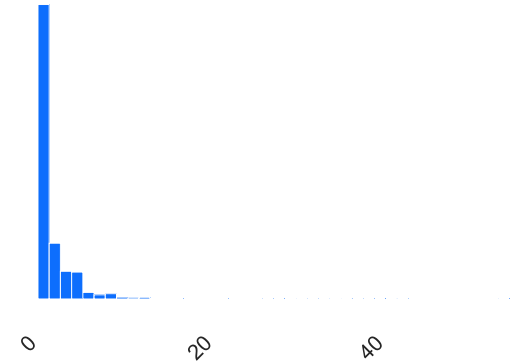
Standard deviation	259.27925
Coefficient of variation (CV)	1.0038494
Kurtosis	20.247938
Mean	258.28501
Median Absolute Deviation (MAD)	94
Skewness	3.2631413
Sum	10638243
Variance	67225.729
Monotonicity	Not monotonic

campaign

Real number (\mathbb{R})

Distinct	42
Distinct (%)	0.1%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	2.5675925

Minimum	1
Maximum	56
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	321.9 KiB



More details

Statistics

Histogram

Common values

Extreme values

Quantile statistics

Minimum	1
5-th percentile	1
Q1	1
median	2
Q3	3
95-th percentile	7
Maximum	56
Range	55
Interquartile range (IQR)	2

Descriptive statistics

Standard deviation	2.7700135
Coefficient of variation (CV)	1.0788369
Kurtosis	36.979795
Mean	2.5675925
Median Absolute Deviation (MAD)	1
Skewness	4.7625067
Sum	105754
Variance	7.672975
Monotonicity	Not monotonic

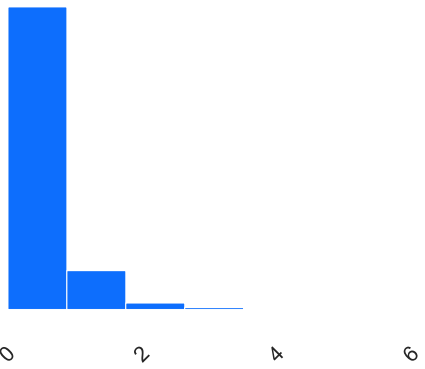
pdays

Real number (\mathbb{R})

High correlation

Distinct	8
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	0.172963

Minimum	0
Maximum	7
Zeros	35563
Zeros (%)	86.3%
Negative	0
Negative (%)	0.0%
Memory size	321.9 KiB



More details

Statistics

Histogram

Common values

Extreme values

Quantile statistics

Minimum	0
5-th percentile	0
Q1	0
median	0
Q3	0
95-th percentile	1
Maximum	7
Range	7
Interquartile range (IQR)	0

Descriptive statistics

Standard deviation	0.49490108
Coefficient of variation (CV)	2.8613119
Kurtosis	20.108816
Mean	0.172963
Median Absolute Deviation (MAD)	0
Skewness	3.8320422
Sum	7124
Variance	0.24492708
Monotonicity	Not monotonic

poutcome

Categorical

- High correlation
- Imbalance

Distinct	3
----------	---

nonexistent	35563
failure	4252

Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	321.9 KiB

success 1373

More details

Overview

Categories

Words

Characters

Length

Max length	11
Median length	11
Mean length	10.453724
Min length	7

Characters and Unicode

Total characters	430568
Distinct characters	13
Distinct categories	1
Distinct scripts	1
Distinct blocks	1

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

Unique

Unique	0
Unique (%)	0.0%

Sample

1st row	nonexistent
2nd row	nonexistent
3rd row	nonexistent
4th row	nonexistent
5th row	nonexistent

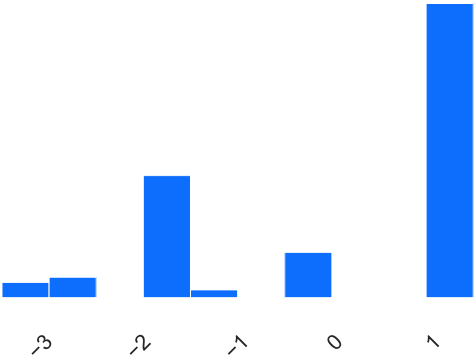
emp.var.rate

Real number (\mathbb{R})

High correlation

Distinct	10
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%

Minimum	-3.4
Maximum	1.4
Zeros	0
Zeros (%)	0.0%
Negative	17191
Negative (%)	41.7%



Mean	0.081885501
------	-------------

Memory size	321.9 KiB
-------------	-----------

More details

Statistics

Histogram

Common values

Extreme values

Quantile statistics

Minimum	-3.4
5-th percentile	-2.9
Q1	-1.8
median	1.1
Q3	1.4
95-th percentile	1.4
Maximum	1.4
Range	4.8
Interquartile range (IQR)	3.2

Descriptive statistics

Standard deviation	1.5709597
Coefficient of variation (CV)	19.184834
Kurtosis	-1.0626315
Mean	0.081885501
Median Absolute Deviation (MAD)	0.3
Skewness	-0.72409555
Sum	3372.7
Variance	2.4679145
Monotonicity	Not monotonic

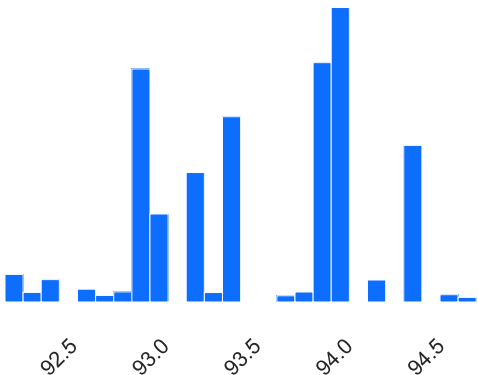
cons.price.idx

Real number (\mathbb{R})

High correlation

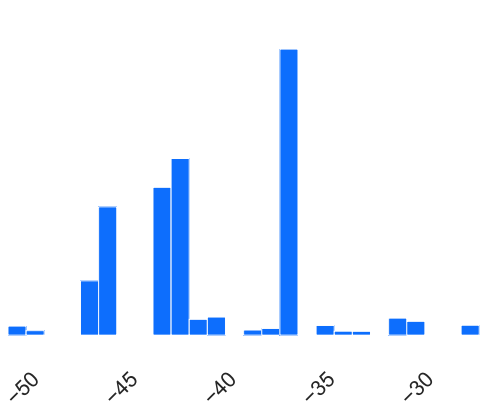
Distinct	26
Distinct (%)	0.1%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	93.575664

Minimum	92.201
Maximum	94.767
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	321.9 KiB



		More details	
Statistics		Histogram	
		Common values	
		Extreme values	
Quantile statistics		Descriptive statistics	
Minimum	92.201	Standard deviation	0.57884005
5-th percentile	92.713	Coefficient of variation (CV)	0.0061857968
Q1	93.075	Kurtosis	-0.82980858
median	93.749	Mean	93.575664
Q3	93.994	Median Absolute Deviation (MAD)	0.38
95-th percentile	94.465	Skewness	-0.23088765
Maximum	94.767	Sum	3854194.5
Range	2.566	Variance	0.3350558
Interquartile range (IQR)	0.919	Monotonicity	Not monotonic

cons.conf.idx			
Real number (\mathbb{R})			
High correlation			
Distinct	26	Minimum	-50.8
Distinct (%)	0.1%	Maximum	-26.9
Missing	0	Zeros	0
Missing (%)	0.0%	Zeros (%)	0.0%
Infinite	0	Negative	41188
Infinite (%)	0.0%	Negative (%)	100.0%
Mean	-40.5026	Memory size	321.9 KiB



A histogram showing the distribution of values for 'cons.conf.idx'. The x-axis represents the value, ranging from -50 to -30. The y-axis represents the frequency. The distribution is unimodal and slightly right-skewed, with a peak frequency of approximately 12 occurring at a value of about -35. There are a few outliers at the lower end of the scale, around -50.

More details

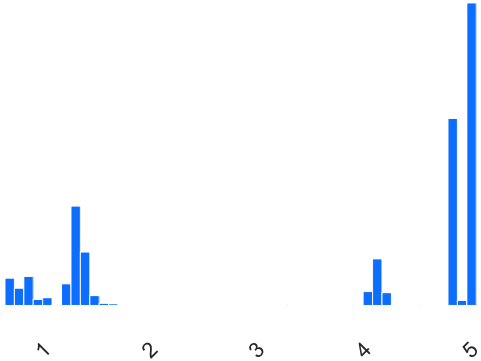
Statistics	Histogram	Common values	Extreme values
Quantile statistics		Descriptive statistics	
Minimum	-50.8	Standard deviation	4.6281979
5-th percentile	-47.1	Coefficient of variation (CV)	-0.11426915
Q1	-42.7	Kurtosis	-0.35855831
median	-41.8	Mean	-40.5026
Q3	-36.4	Median Absolute Deviation (MAD)	4.4
95-th percentile	-33.6	Skewness	0.30317986
Maximum	-26.9	Sum	-1668221.1
Range	23.9	Variance	21.420215
Interquartile range (IQR)	6.3	Monotonicity	Not monotonic

euribor3m

Real number (\mathbb{R})

High correlation

Distinct	316	Minimum	0.634
Distinct (%)	0.8%	Maximum	5.045
Missing	0	Zeros	0
Missing (%)	0.0%	Zeros (%)	0.0%
Infinite	0	Negative	0
Infinite (%)	0.0%	Negative (%)	0.0%
Mean	3.6212908	Memory size	321.9 KiB



More details

Quantile statistics

Minimum	0.634
5-th percentile	0.797
Q1	1.344
median	4.857
Q3	4.961
95-th percentile	4.966
Maximum	5.045
Range	4.411
Interquartile range (IQR)	3.617

Descriptive statistics

Standard deviation	1.7344474
Coefficient of variation (CV)	0.47895833
Kurtosis	-1.4068026
Mean	3.6212908
Median Absolute Deviation (MAD)	0.108
Skewness	-0.70918796
Sum	149153.73
Variance	3.0083078
Monotonicity	Not monotonic

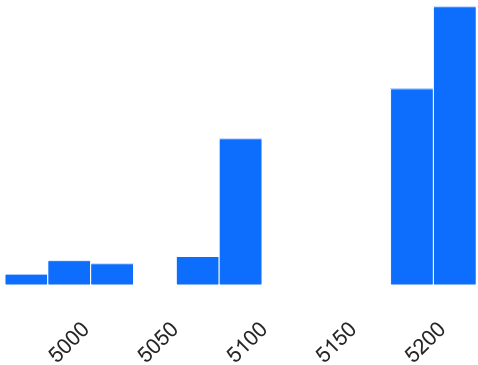
nr.employed

Real number (\mathbb{R})

High correlation

Distinct	11
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	5167.0359

Minimum	4963.6
Maximum	5228.1
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	321.9 KiB



More details

Statistics

Histogram

Common values

Extreme values

Quantile statistics

Descriptive statistics

Minimum	4963.6	Standard deviation	72.251528
5-th percentile	5017.5	Coefficient of variation (CV)	0.013983167
Q1	5099.1	Kurtosis	-0.0037603757
median	5191	Mean	5167.0359
Q3	5228.1	Median Absolute Deviation (MAD)	37.1
95-th percentile	5228.1	Skewness	-1.0442624
Maximum	5228.1	Sum	2.1281988×10^8
Range	264.5	Variance	5220.2833
Interquartile range (IQR)	129	Monotonicity	Not monotonic

y

Boolean

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	40.3 KiB

False36548

True4640

More details

Common Values (Table)

Common Values (Plot)

Value	Count	Frequency (%)
False	36548	88.7%
True	4640	11.3%

Interactions

age

duration

campaign

pdays

previous

emp.var.rate

cons.price.idx

cons.conf.idx

euribor3m

nr.employed

nr.employed

age

duration

campaign

pdays

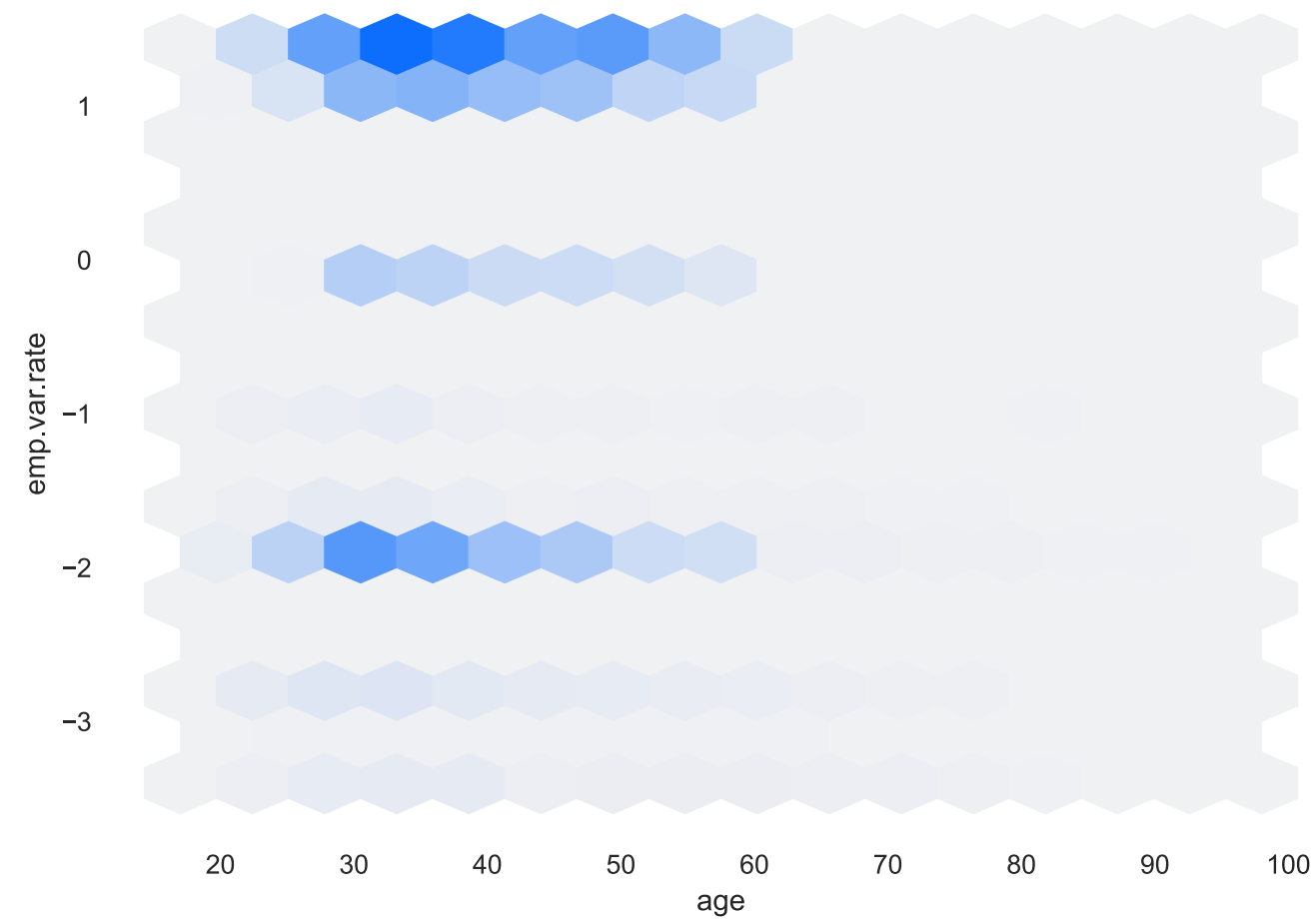
previous

emp.var.rate

cons.price.idx

cons.conf.idx

euribor3m

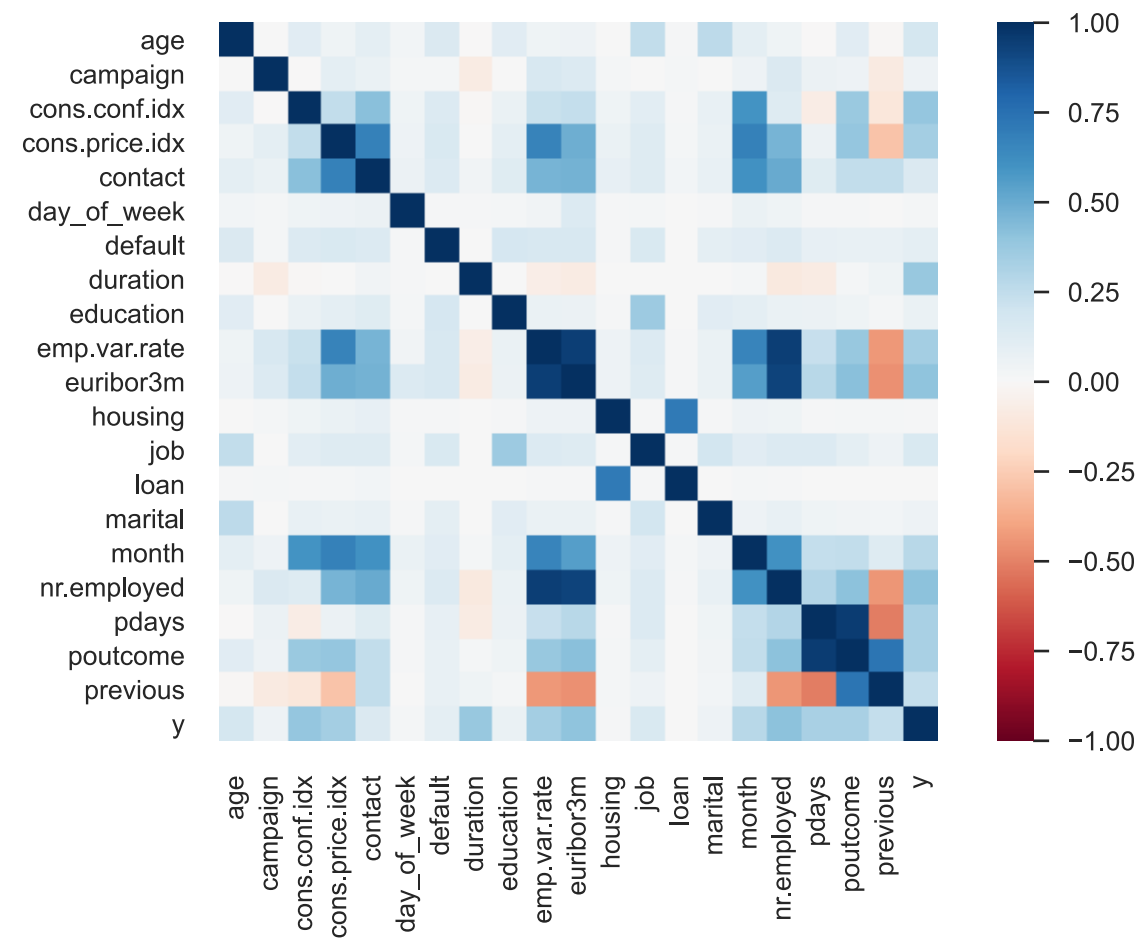


Correlations

Auto

Heatmap

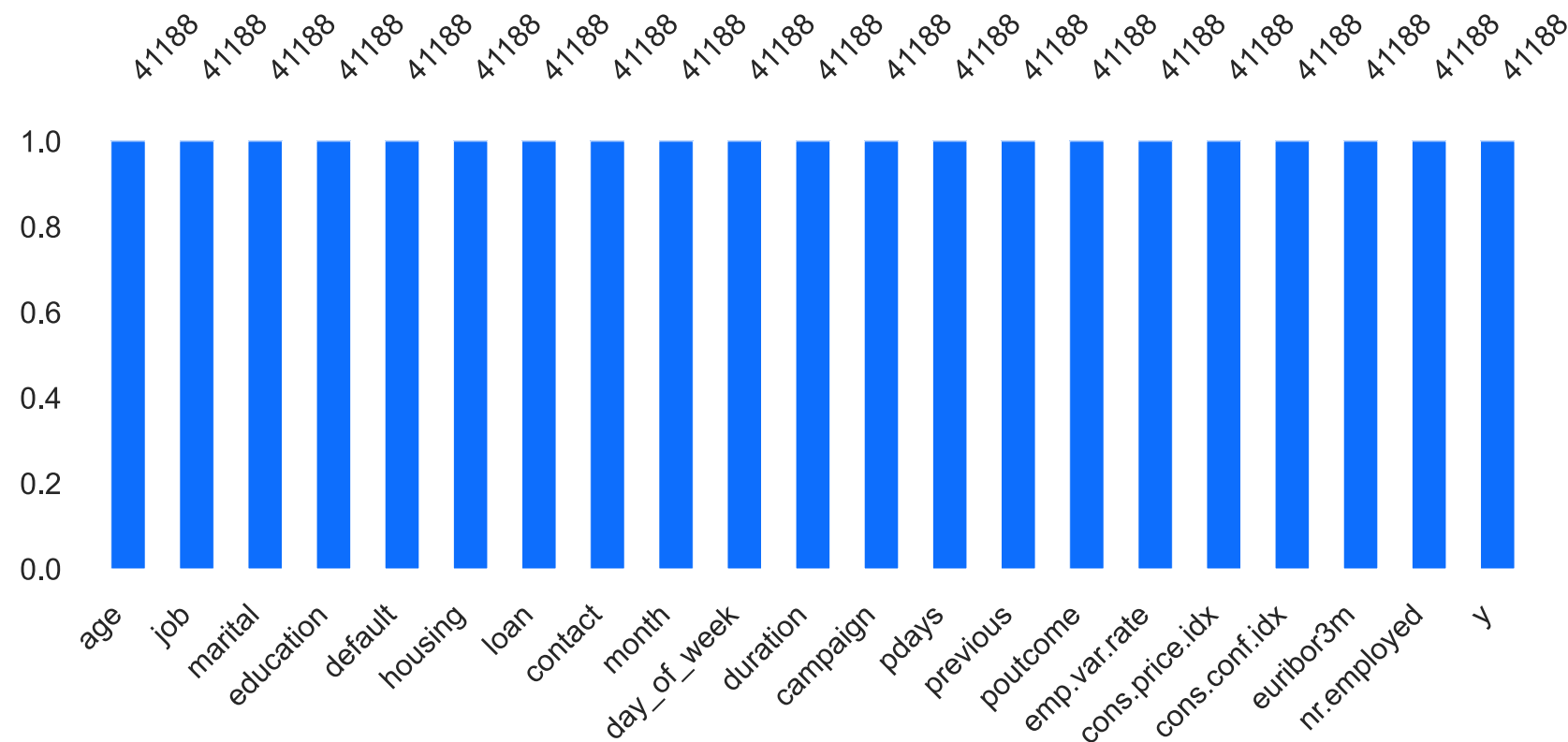
Table



Missing values

Count

[Matrix](#)



A simple visualization of nullity by column.

Sample

First rows

Last rows

	age	job	marital	education	default	housing	loan	contact	month	day_of_week	duration	campaign	pdays	previous	poutcome
41178	62	retired	married	university.degree	no	no	no	cellular	nov	thu	483	2	6	3	success
41179	64	retired	divorced	professional.course	no	yes	no	cellular	nov	fri	151	3	999	0	nonexist
41180	36	admin.	married	university.degree	no	no	no	cellular	nov	fri	254	2	999	0	nonexist
41181	37	admin.	married	university.degree	no	yes	no	cellular	nov	fri	281	1	999	0	nonexist
41182	29	unemployed	single	basic.4y	no	yes	no	cellular	nov	fri	112	1	9	1	success
41183	73	retired	married	professional.course	no	yes	no	cellular	nov	fri	334	1	999	0	nonexist
41184	46	blue-collar	married	professional.course	no	no	no	cellular	nov	fri	383	1	999	0	nonexist
41185	56	retired	married	university.degree	no	yes	no	cellular	nov	fri	189	2	999	0	nonexist

	age	job	marital	education	default	housing	loan	contact	month	day_of_week	duration	campaign	pdays	previous	poutcome
41186	44	technician	married	professional.course	no	no	no	cellular	nov	fri	442	1	999	0	nonexistent
41187	74	retired	married	professional.course	no	yes	no	cellular	nov	fri	239	3	999	1	failure

Duplicate rows

Most frequently occurring

	age	job	marital	education	default	housing	loan	contact	month	day_of_week	duration	campaign	pdays	previous	poutcome
0	24	services	single	high.school	no	yes	no	cellular	apr	tue	114	1	999	0	nonexistent
1	27	technician	single	professional.course	no	no	no	cellular	jul	mon	331	2	999	0	nonexistent
2	32	technician	single	professional.course	no	yes	no	cellular	jul	thu	128	1	999	0	nonexistent
3	35	admin.	married	university.degree	no	yes	no	cellular	may	fri	348	4	999	0	nonexistent
4	36	retired	married	unknown	no	no	no	telephone	jul	thu	88	1	999	0	nonexistent
5	39	admin.	married	university.degree	no	no	no	cellular	nov	tue	123	2	999	0	nonexistent
6	39	blue-collar	married	basic.6y	no	no	no	telephone	may	thu	124	1	999	0	nonexistent
7	41	technician	married	professional.course	no	yes	no	cellular	aug	tue	127	1	999	0	nonexistent
8	45	admin.	married	university.degree	no	no	no	cellular	jul	thu	252	1	999	0	nonexistent
9	47	technician	divorced	high.school	no	yes	no	cellular	jul	thu	43	3	999	0	nonexistent

