

Overview

Brought to you by [YData](#)

Overview

Alerts 17

Reproduction

Dataset statistics

Number of variables	14
Number of observations	889
Missing cells	0
Missing cells (%)	0.0%
Duplicate rows	58
Duplicate rows (%)	6.5%
Total size in memory	86.1 KiB
Average record size in memory	99.1 B

Variable types

Categorical	7
Numeric	4
Boolean	3

Variables

Select Columns

survived

Categorical

High correlation

Distinct	2	0	549
Distinct (%)	0.2%	1	340
Missing	0		
Missing (%)	0.0%		
Memory size	13.9 KiB		

More details

pclass

Categorical

High correlation

Distinct	3	3	491
Distinct (%)	0.3%	1	214
Missing	0	2	184
Missing (%)	0.0%		
Memory size	13.9 KiB		

More details

sex

Categorical

High correlation

Distinct	2	male	577
Distinct (%)	0.2%	female	312
Missing	0		

Missing (%)0.0%

Memory size13.9 KiB

More details

age

Real number (ℝ)

High correlation

Distinct88

Distinct (%)9.9%

Missing0

Missing (%)0.0%

Infinite0

Infinite (%)0.0%

Mean29.315152

Minimum0.42

Maximum80

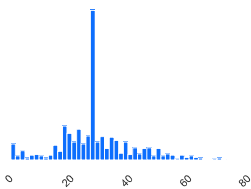
Zeros0

Zeros (%)0.0%

Negative0

Negative (%)0.0%

Memory size13.9 KiB



◀ ▶

◀ ▶

More details

sibsp

Real number (ℝ)

High correlationZeros

Distinct7

Distinct (%)0.8%

Missing0

Missing (%)0.0%

Infinite0

Infinite (%)0.0%

Mean0.52418448

Minimum0

Maximum8

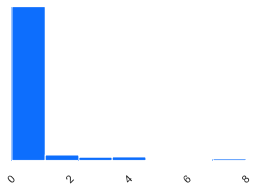
Zeros606

Zeros (%)68.2%

Negative0

Negative (%)0.0%

Memory size13.9 KiB



A histogram showing the distribution of survival status. The x-axis represents survival status (0 to 6) and the y-axis represents frequency (0 to 6). The distribution is highly skewed towards 0, with a frequency of 6.

More details

parch

Real number (\mathbb{R})

High correlationZeros

Distinct7

Distinct (%)0.8%

Missing0

Missing (%)0.0%

Infinite0

Infinite (%)0.0%

Mean0.38245219

Minimum0

Maximum6

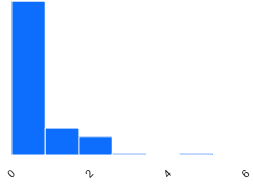
Zeros676

Zeros (%)76.0%

Negative0

Negative (%)0.0%

Memory size13.9 KiB



A histogram showing the distribution of the parch variable. The x-axis represents parch values (0 to 6) and the y-axis represents frequency (0 to 6). The distribution is highly skewed towards 0, with a frequency of 6.

More details

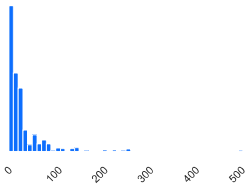
More details

fare

Real number (ℝ)

Zeros

Distinct	247	Minimum	0
Distinct (%)	27.8%	Maximum	512.32
Missing	0	Zeros	15
Missing (%)	0.0%	Zeros (%)	1.7%
Infinite	0	Negative	0
Infinite (%)	0.0%	Negative (%)	0.0%
Mean	32.096681	Memory size	13.9 KiB



More details

embarked

Categorical

High correlation

Distinct	3	S	644
Distinct (%)	0.3%	C	168
Missing	0	Q	77
Missing (%)	0.0%		

Memory size 13.9 KiB

More details

class

Categorical

High correlation

Distinct 3

Distinct (%) 0.3%

Missing 0

Missing (%) 0.0%

Memory size 7.9 KiB

Third 491
First 214
Second 184

More details

who

Categorical

High correlation

Distinct 3

Distinct (%) 0.3%

Missing 0

Missing (%) 0.0%

Memory size 13.9 KiB

man 537
woman 269
child 83

More details

adult_male

Boolean

High correlation

Distinct	2	True	537
Distinct (%)	0.2%	False	352
Missing	0		
Missing (%)	0.0%		
Memory size	7.8 KiB		

More details

embark_town

Categorical

High correlation

Distinct	3	Southam...	644
Distinct (%)	0.3%	Cherbou...	168
Missing	0	Queenst...	77
Missing (%)	0.0%		
Memory size	13.9 KiB		

More details

alive

Boolean

High correlation

Distinct	2	False	549
Distinct (%)	0.2%	True	340
Missing	0		
Missing (%)	0.0%		
Memory size	7.8 KiB		

More details

alone

Boolean

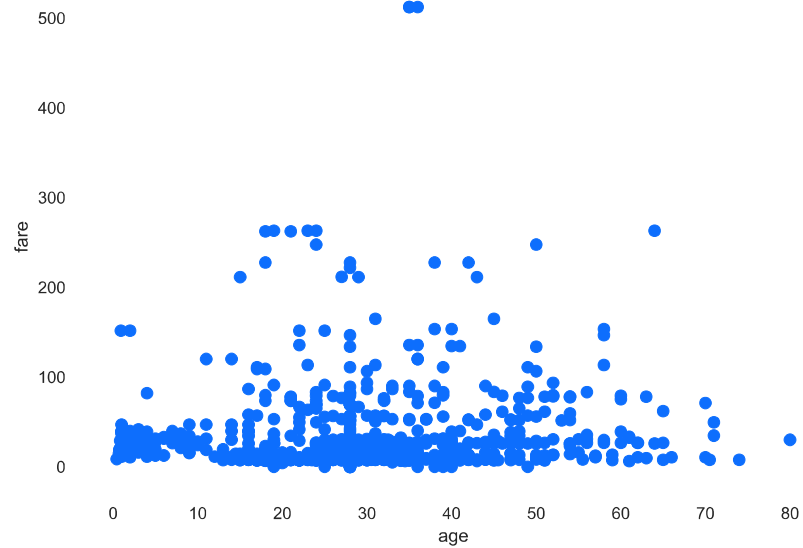
High correlation

Distinct	2	True	535
Distinct (%)	0.2%	False	354
Missing	0		
Missing (%)	0.0%		
Memory size	7.8 KiB		

More details

Interactions

age	sibsp	parch	fare
fare	age	sibsp	parch

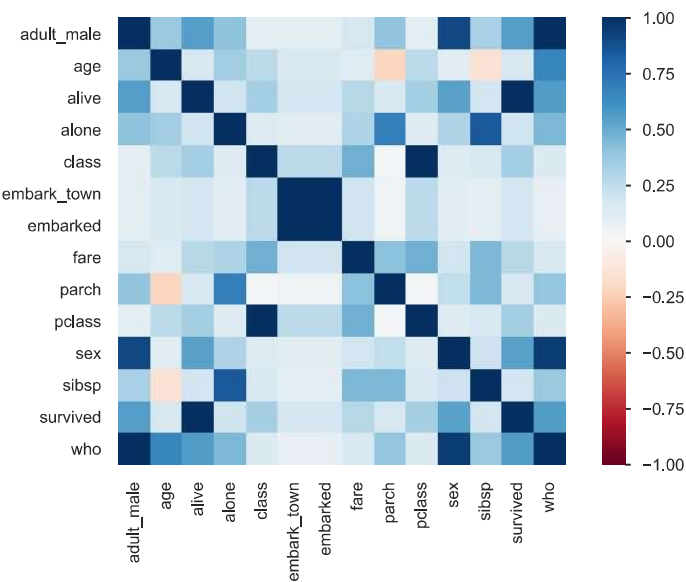


Correlations

Auto

Heatmap

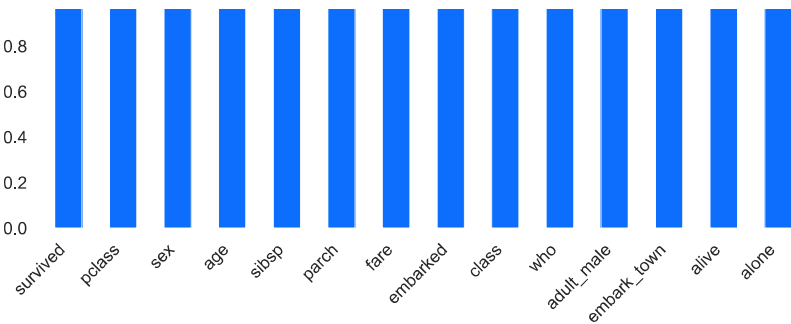
[Table](#)



Missing values

Count [Matrix](#)

Titanic EDA Report



A simple visualization of nullity by column.

Sample

First rows

Last rows

	survived	pclass	sex	age	sibsp	parch	fare	en
0	0	3	male	22.0	1	0	7.2500	S
1	1	1	female	38.0	1	0	71.2833	C
2	1	3	female	26.0	0	0	7.9250	S
3	1	1	female	35.0	1	0	53.1000	S
4	0	3	male	35.0	0	0	8.0500	S
5	0	3	male	28.0	0	0	8.4583	Q

	survived	pclass	sex	age	sibsp	parch	fare	en
6	0	1	male	54.0	0	0	51.8625	S
7	0	3	male	2.0	3	1	21.0750	S
8	1	3	female	27.0	0	2	11.1333	S
9	1	2	female	14.0	1	0	30.0708	C

Duplicate rows

Most frequently occurring

	survived	pclass	sex	age	sibsp	parch	fare	en
30	0	3	male	28.0	0	0	7.8958	S
31	0	3	male	28.0	0	0	8.0500	S
28	0	3	male	28.0	0	0	7.7500	C
51	1	3	female	28.0	0	0	7.7500	C
7	0	2	male	28.0	0	0	0.0000	S
25	0	3	male	28.0	0	0	7.2250	C
26	0	3	male	28.0	0	0	7.2292	C
37	0	3	male	28.0	8	2	69.5500	S
4	0	2	male	23.0	0	0	13.0000	S
5	0	2	male	25.0	0	0	13.0000	S

