

# Overview

## Dataset statistics

Number of variables	11
Number of observations	20433
Missing cells	0
Missing cells (%)	0.0%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	1.7 MiB
Average record size in memory	88.0 B

## Variable types

Numeric	10
Categorical	1

# Variables

df\_index

Real number ( $\mathbb{R}_{\geq 0}$ )

HIGH

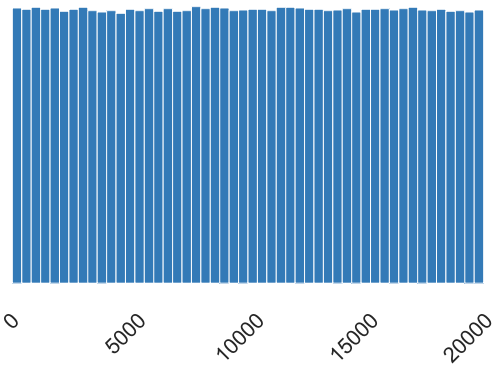
CORRELATION

UNIFORM

UNIQUE

Distinct	20433
Distinct (%)	100.0%
Missing	0
Missing (%)	0.0%

Mean	10316.17609
Minimum	0
Maximum	20639
Zeros	1
Zeros (%)	< 0.1%
Negative	0
Negative (%)	0.0%
Memory size	159.8 KiB



Toggle details

longitude

Real number ( $\mathbb{R}$ )

HIGH

CORRELATION

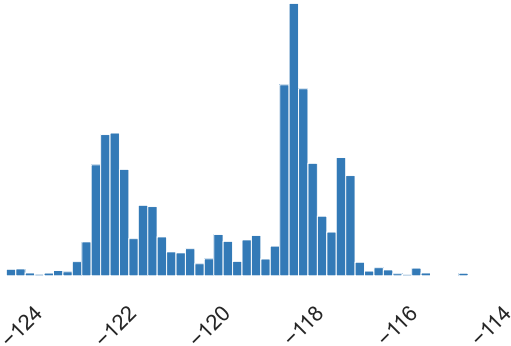
HIGH CORRELATION

HIGH CORRELATION

HIGH CORRELATION

Distinct	844
Distinct (%)	4.1%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	-119.5706886
Minimum	-124.35
Maximum	-114.31
Zeros	0

Negative	20455
Negative (%)	100.0%
Memory size	159.8 KiB



Toggle details

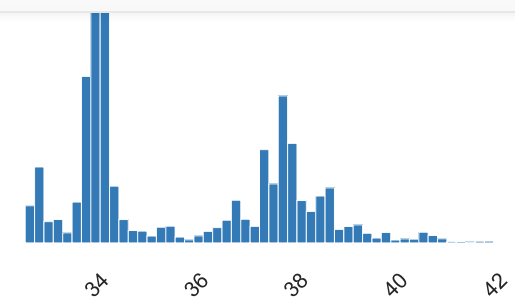
latitude

Real number ( $\mathbb{R}_{\geq 0}$ )

HIGH

CORRELATION  
HIGH CORRELATION  
HIGH CORRELATION  
HIGH CORRELATION

Distinct	861
Distinct (%)	4.2%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	35.63322126
Minimum	32.54
Maximum	41.95
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	159.8 KiB

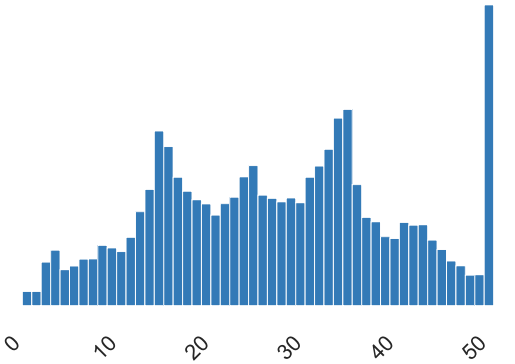


Toggle details

housing\_median\_age

Real number ( $\mathbb{R}_{\geq 0}$ )

Distinct	52
Distinct (%)	0.3%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	28.63309353
Minimum	1
Maximum	52
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	159.8 KiB



Toggle details

total\_rooms

Real number ( $\mathbb{R}_{\geq 0}$ )

HIGH

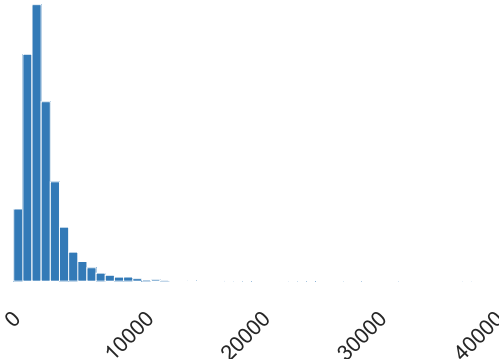
CORRELATION

HIGH CORRELATION

HIGH CORRELATION

HIGH CORRELATION

Distinct	5911
Distinct (%)	28.9%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	2636.504233
Minimum	2
Maximum	39320
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	159.8 KiB



Toggle details

total\_bedrooms

Real number ( $\mathbb{R}_{\geq 0}$ )

HIGH

CORRELATION

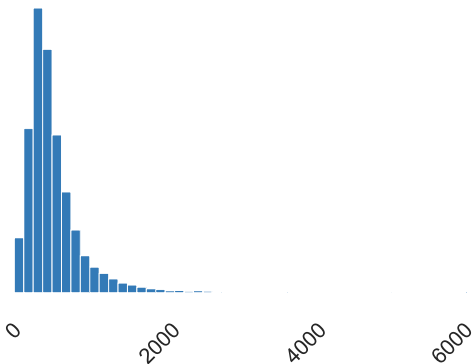
HIGH CORRELATION

HIGH CORRELATION

HIGH CORRELATION



Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	537.8705525
Minimum	1
Maximum	6445
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	159.8 KiB



Toggle details

population

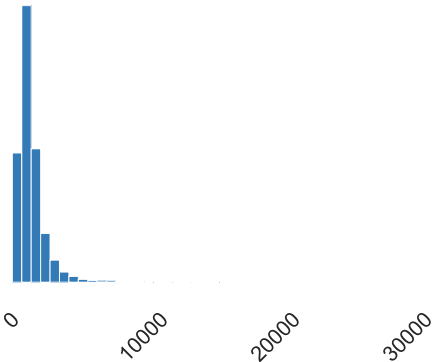
Real number ( $\mathbb{R}_{\geq 0}$ )

HIGH

- CORRELATION
- HIGH CORRELATION
- HIGH CORRELATION
- HIGH CORRELATION

Distinct	3879
Distinct (%)	19.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%

Minimum	3
Maximum	35682
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	159.8 KiB



Toggle details

households

Real number ( $\mathbb{R}_{\geq 0}$ )

HIGH

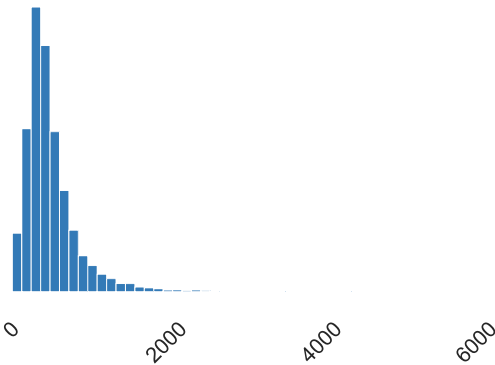
CORRELATION  
HIGH CORRELATION  
HIGH CORRELATION  
HIGH CORRELATION

Distinct	1809
Distinct (%)	8.9%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	499.4334655

Minimum	1
Maximum	6082
Zeros	0
Zeros (%)	0.0%
Negative	0

Memory size

159.8 KiB



Toggle details

median\_income

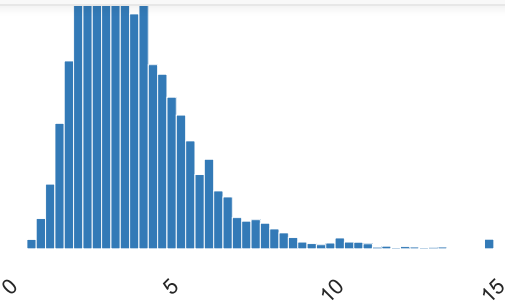
Real number ( $\mathbb{R}_{\geq 0}$ )

HIGH

CORRELATION  
HIGH CORRELATION  
HIGH CORRELATION

Distinct	12825
Distinct (%)	62.8%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	3.871161601
Minimum	0.4999
Maximum	15.0001
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	159.8 KiB





Toggle details

median\_house\_value

Real number ( $\mathbb{R}_{\geq 0}$ )

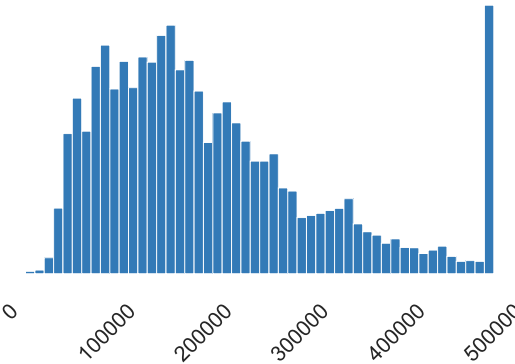
HIGH

CORRELATION

HIGH CORRELATION

HIGH CORRELATION

Distinct	3833
Distinct (%)	18.8%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	206864.4132
Minimum	14999
Maximum	500001
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	159.8 KiB



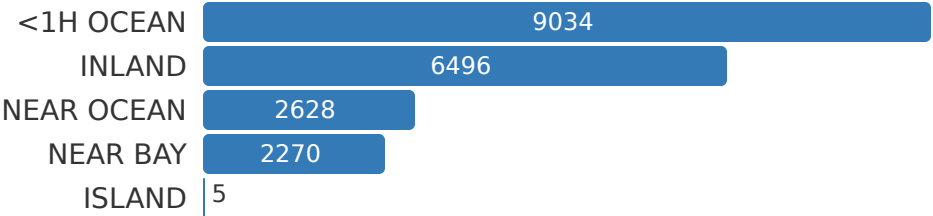
ocean\_proximity

Categorical

HIGH

CORRELATION

Distinct	5
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	159.8 KiB



Toggle details

Interactions

df\_index

longitude

latitude

housing\_median\_age

total\_rooms

total\_bedrooms

population

households

median\_income

median\_house\_value

median\_house\_value

df\_index

longitude

latitude

housing\_median\_age

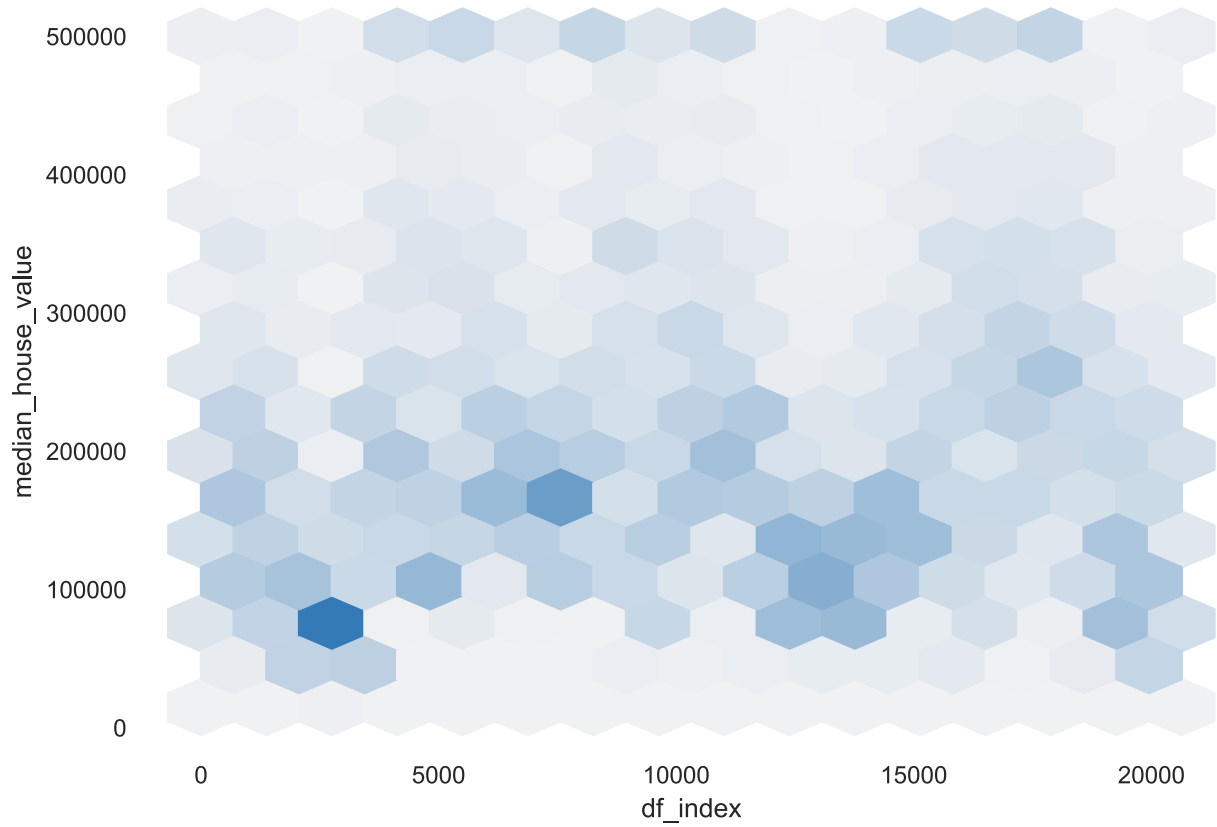
total\_rooms

total\_bedrooms

population

households

median\_income



# Correlations

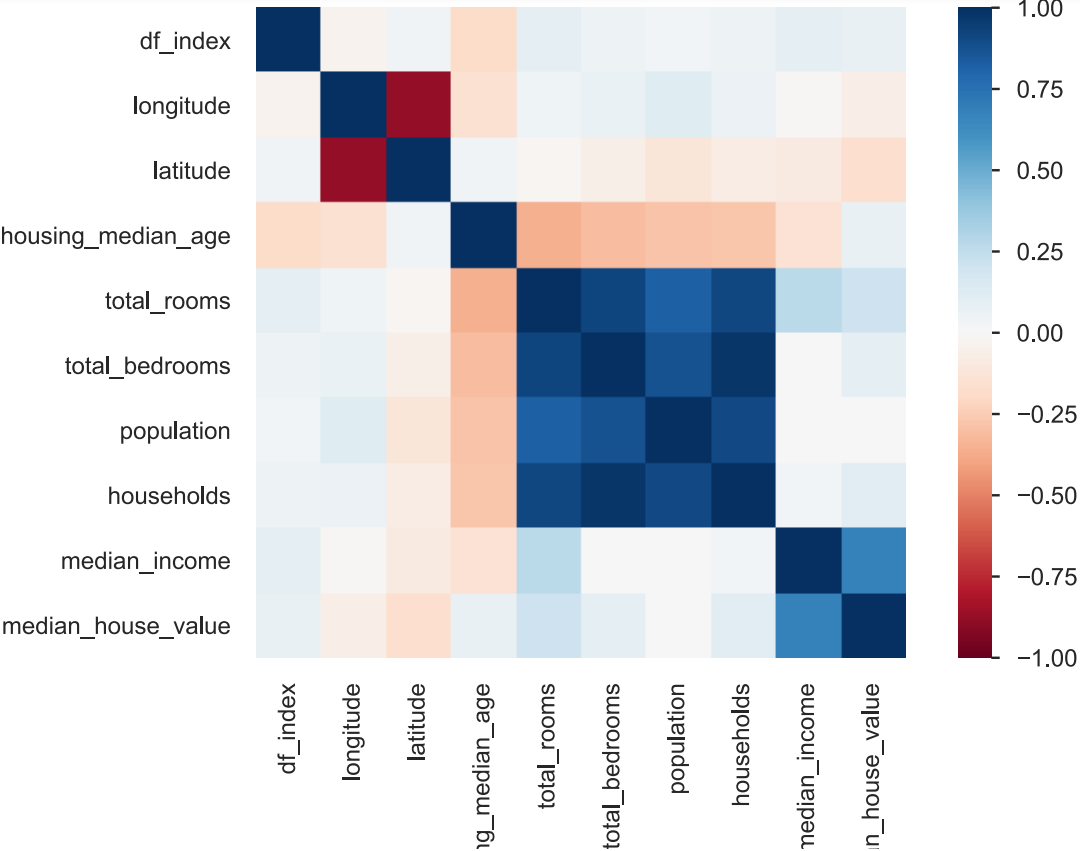
Spearman's  $\rho$

Pearson's  $r$

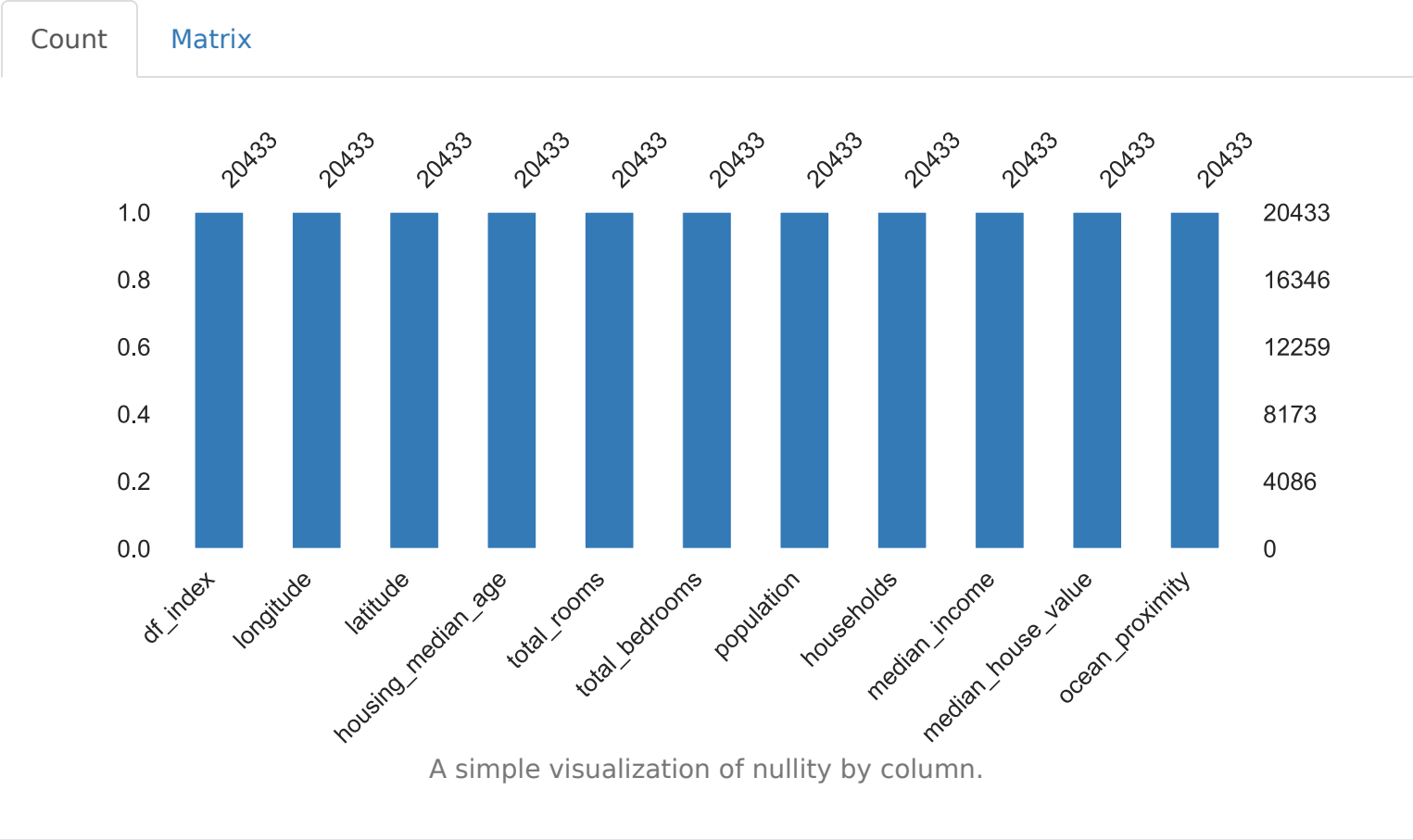
Kendall's  $\tau$

Phik ( $\phi_k$ )

Toggle correlation descriptions



# Missing values



First rows

	df_index	longitude	latitude	housing_median_age	total_rooms	total_bedrooms	population
0	0	-122.23	37.88	41.0	880.0	129.0	3200
1	1	-122.22	37.86	21.0	7099.0	1106.0	2400
2	2	-122.24	37.85	52.0	1467.0	190.0	4900
3	3	-122.25	37.85	52.0	1274.0	235.0	5500
4	4	-122.25	37.85	52.0	1627.0	280.0	5600
5	5	-122.25	37.85	52.0	919.0	213.0	4100
6	6	-122.25	37.84	52.0	2535.0	489.0	1000
7	7	-122.25	37.84	52.0	3104.0	687.0	1100
8	8	-122.26	37.84	42.0	2555.0	665.0	1200
9	9	-122.25	37.84	52.0	3549.0	707.0	1500

Last rows

	df_index	longitude	latitude	housing_median_age	total_rooms	total_bedrooms	population
20423	20630	-121.32	39.29	11.0	2640.0	505.0	1500
20424	20631	-121.40	39.33	15.0	2655.0	493.0	1500
20425	20632	-121.45	39.26	15.0	2319.0	416.0	1500
20426	20633	-121.53	39.19	27.0	2080.0	412.0	1500
20427	20634	-121.56	39.27	28.0	2332.0	395.0	1500
20428	20635	-121.09	39.48	25.0	1665.0	374.0	1500
20429	20636	-121.21	39.49	18.0	697.0	150.0	1500
20430	20637	-121.22	39.43	17.0	2254.0	485.0	1500
20431	20638	-121.32	39.43	18.0	1860.0	409.0	1500
20432	20639	-121.24	39.37	16.0	2785.0	616.0	1500



Report generated with [pandas-profiling](#).