



Overview

Brought to you by [YData](#)

- Overview
- Alerts11
- Reproduction

Dataset statistics

Number of variables	15
Number of observations	48842
Missing cells	2203
Missing cells (%)	0.3%
Duplicate rows	28
Duplicate rows (%)	0.1%
Total size in memory	29.8 MiB
Average record size in memory	639.4 B

Variable types

Numeric	6
Categorical	9

Variables

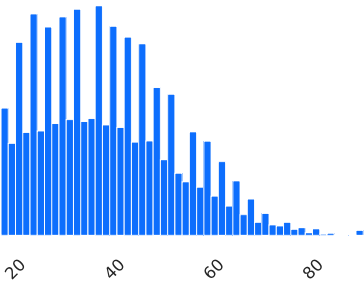
Select Columns

age

Real number (\mathbb{R})

Distinct	74
Distinct (%)	0.2%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	38.643585
Minimum	17
Maximum	90
Zeros	0

Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	381.7 KiB



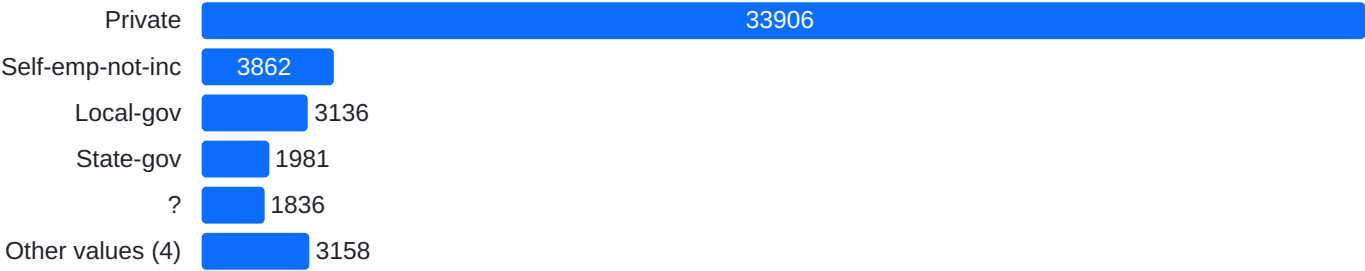
More details

workclass

Categorical

Missing

Distinct	9
Distinct (%)	< 0.1%
Missing	963
Missing (%)	2.0%
Memory size	3.0 MiB



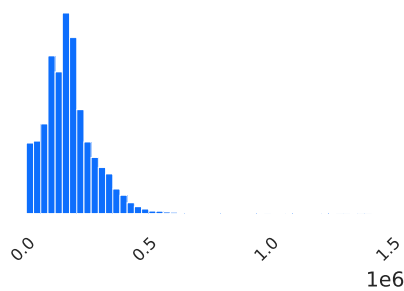
More details

fnlwgt

Real number (\mathbb{R})

Distinct	28523
Distinct (%)	58.4%
Missing	0
Missing (%)	0.0%
Infinite	0

Infinite (%)	0.0%
Mean	189664.13
Minimum	12285
Maximum	1490400
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	381.7 KiB



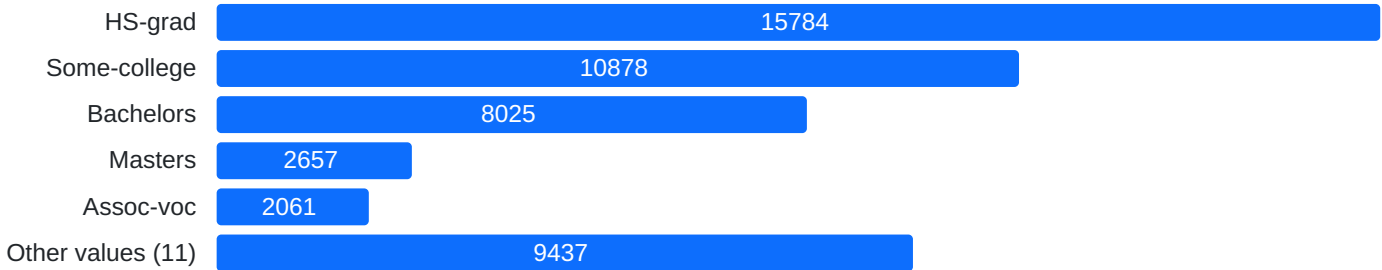
[More details](#)

education

Categorical

High correlation

Distinct	16
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	3.0 MiB



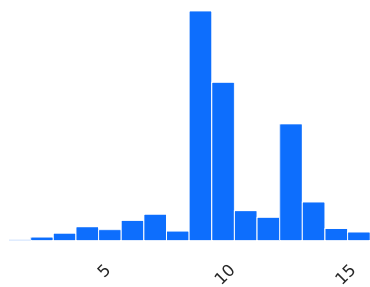
[More details](#)

education-num

Real number (\mathbb{R})

High correlation

Distinct	16
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	10.078089
Minimum	1
Maximum	16
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	381.7 KiB

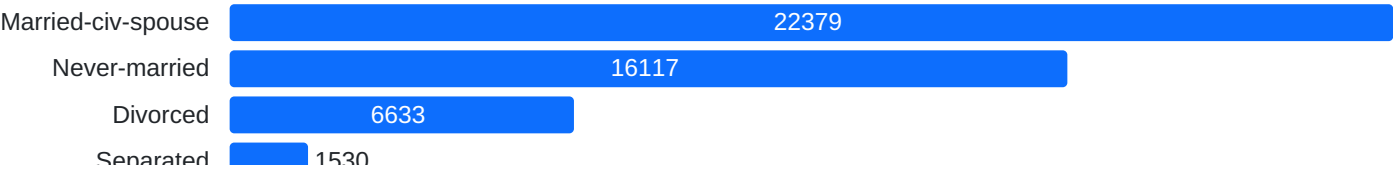


More details

marital-status

Categorical

Distinct	7
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	3.3 MiB



Separated 1000
Widowed 1518
Other values (2) 665

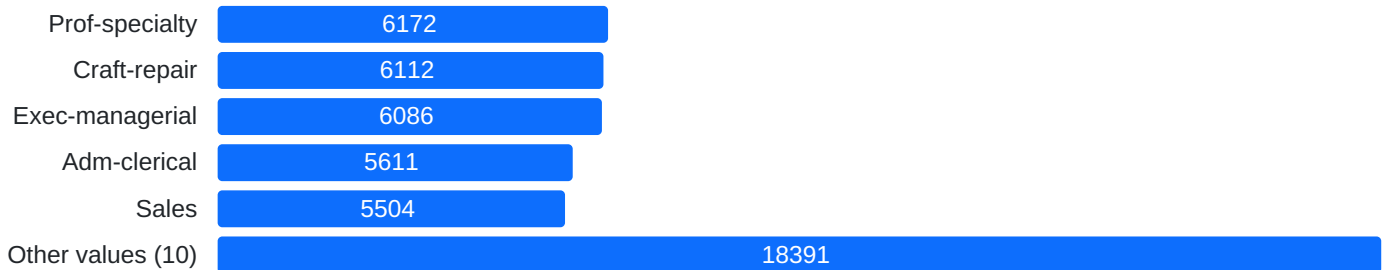
More details

occupation

Categorical

Missing

Distinct	15
Distinct (%)	< 0.1%
Missing	966
Missing (%)	2.0%
Memory size	3.2 MiB



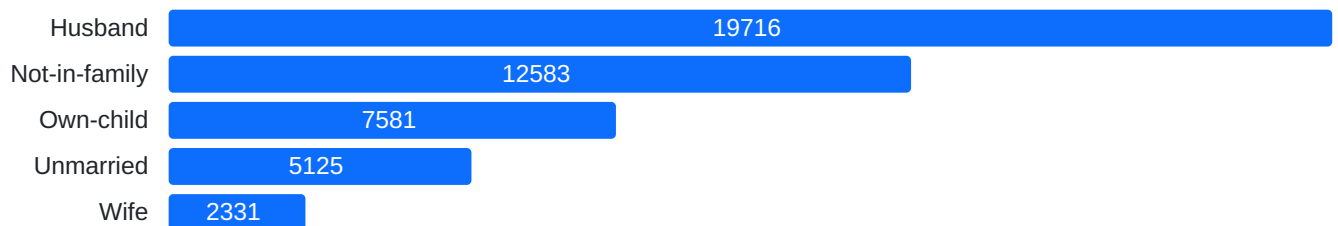
More details

relationship

Categorical

High correlation

Distinct	6
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	3.1 MiB



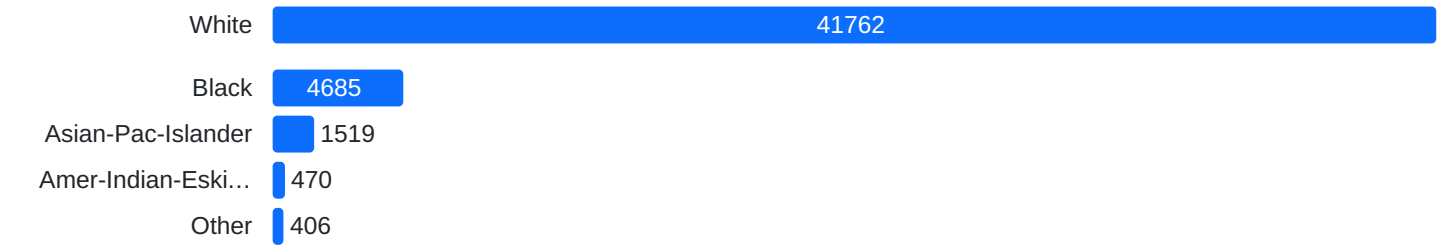
More details

race

Categorical

Imbalance

Distinct	5
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	2.9 MiB



More details

sex

Categorical

High correlation

Distinct	2
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	2.9 MiB



More details

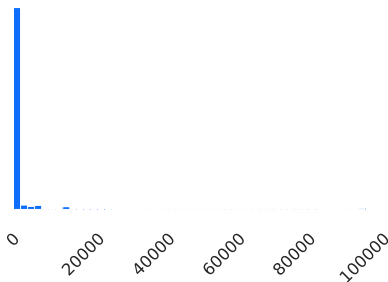
capital-gain

Real number (\mathbb{R})

Zeros

Distinct	123
Distinct (%)	0.3%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%

Mean	1079.0676
Minimum	0
Maximum	99999
Zeros	44807
Zeros (%)	91.7%
Negative	0
Negative (%)	0.0%
Memory size	381.7 KiB



More details

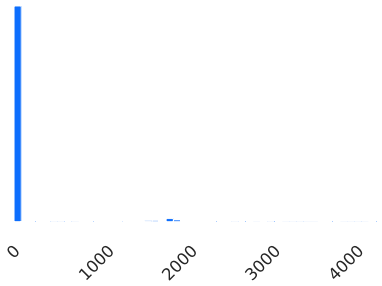
capital-loss

Real number (\mathbb{R})

Zeros

Distinct	99
Distinct (%)	0.2%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	87.502314

Minimum	0
Maximum	4356
Zeros	46560
Zeros (%)	95.3%
Negative	0
Negative (%)	0.0%
Memory size	381.7 KiB



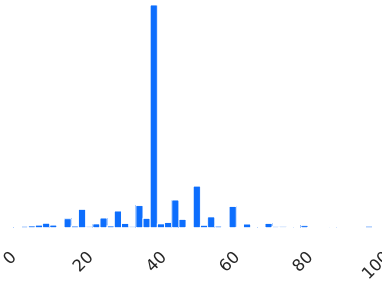
[More details](#)

hours-per-week

Real number (\mathbb{R})

Distinct	96
Distinct (%)	0.2%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Mean	40.422382

Minimum	1
Maximum	99
Zeros	0
Zeros (%)	0.0%
Negative	0
Negative (%)	0.0%
Memory size	381.7 KiB



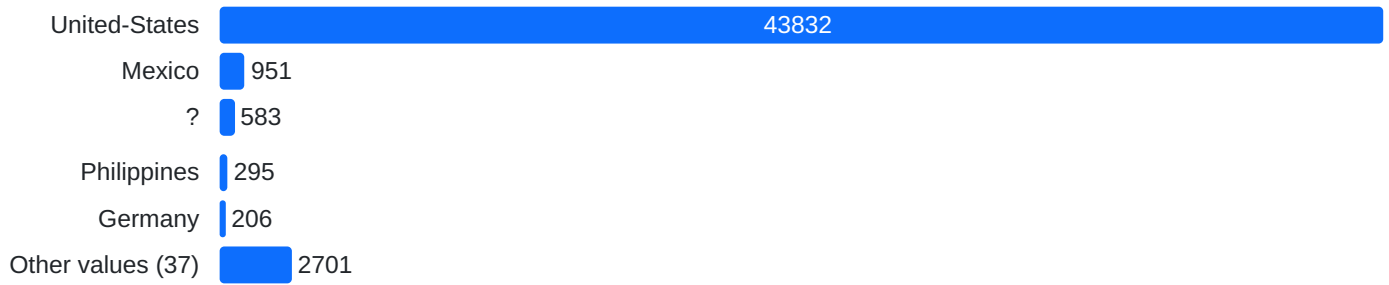
[More details](#)

native-country

Categorical

Imbalance

Distinct	42
Distinct (%)	0.1%
Missing	274
Missing (%)	0.6%
Memory size	3.2 MiB

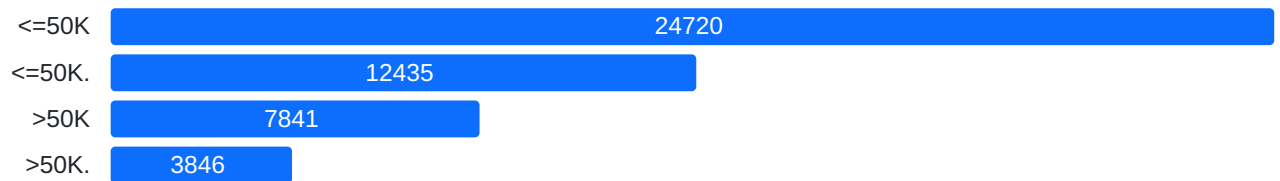


More details

income

Categorical

Distinct	4
Distinct (%)	< 0.1%
Missing	0
Missing (%)	0.0%
Memory size	2.9 MiB



More details

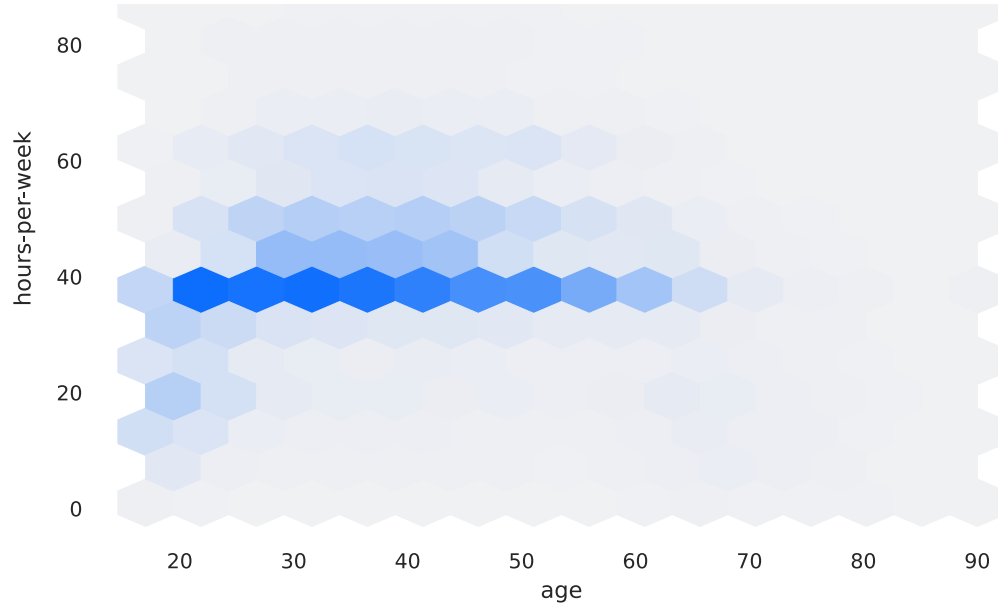
Interactions

age fnlwgt education-num capital-gain capital-loss hours-per-week

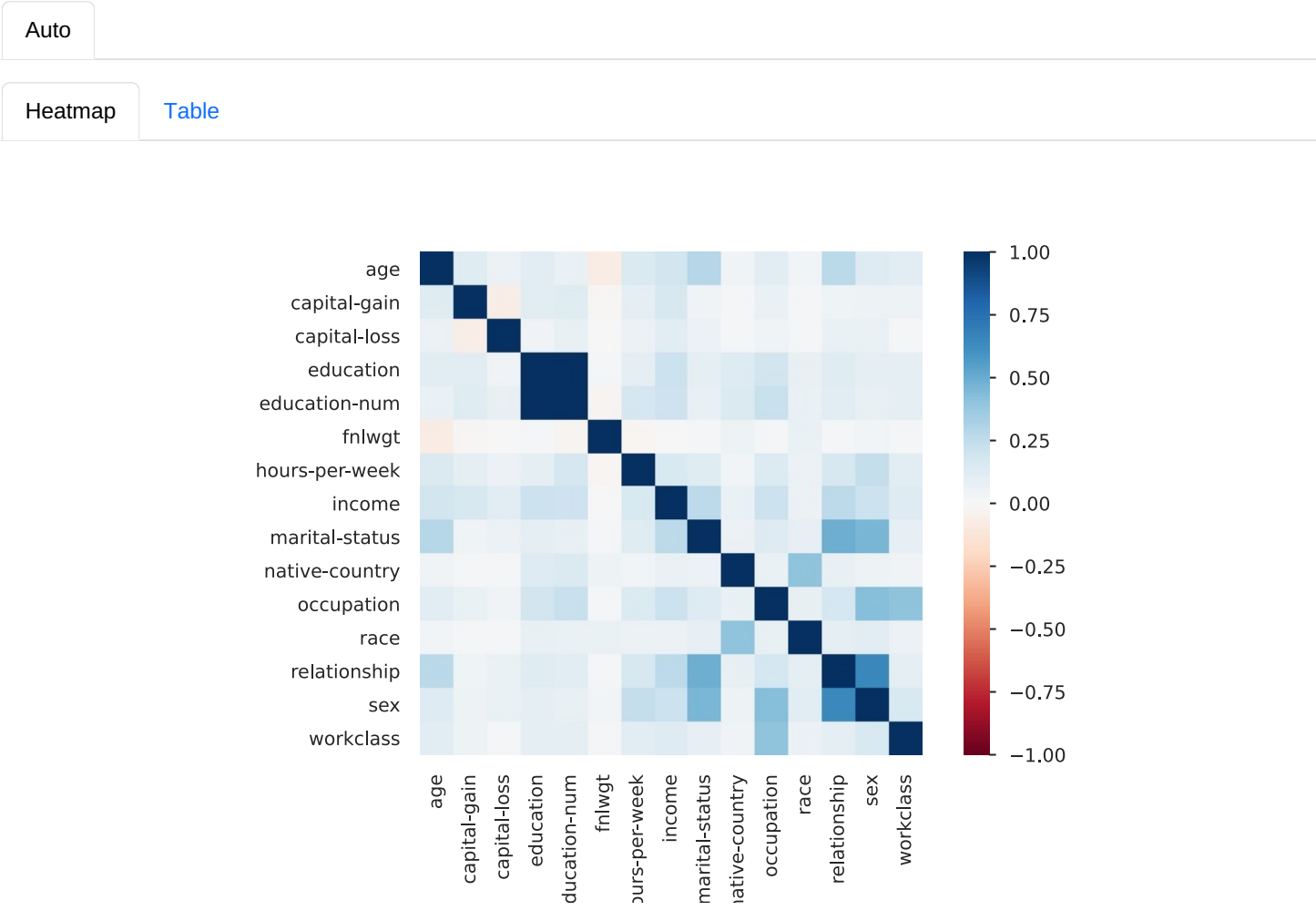
hours-per-week age fnlwgt education-num capital-gain capital-loss

100

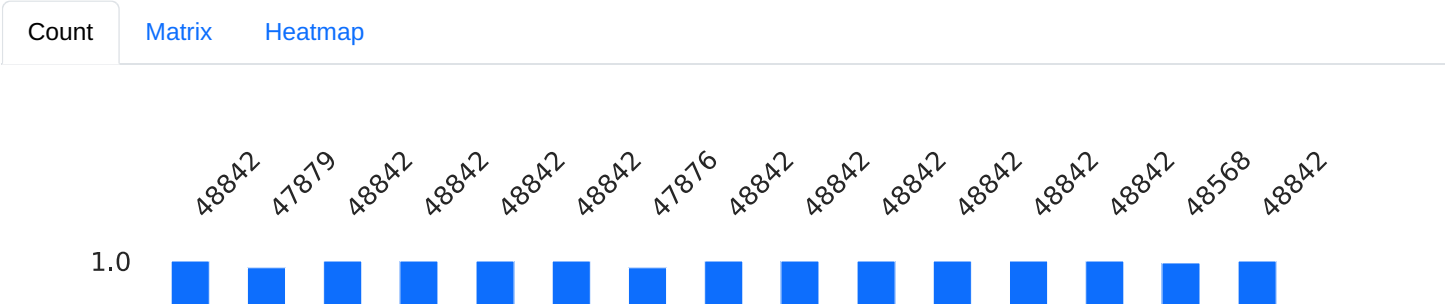


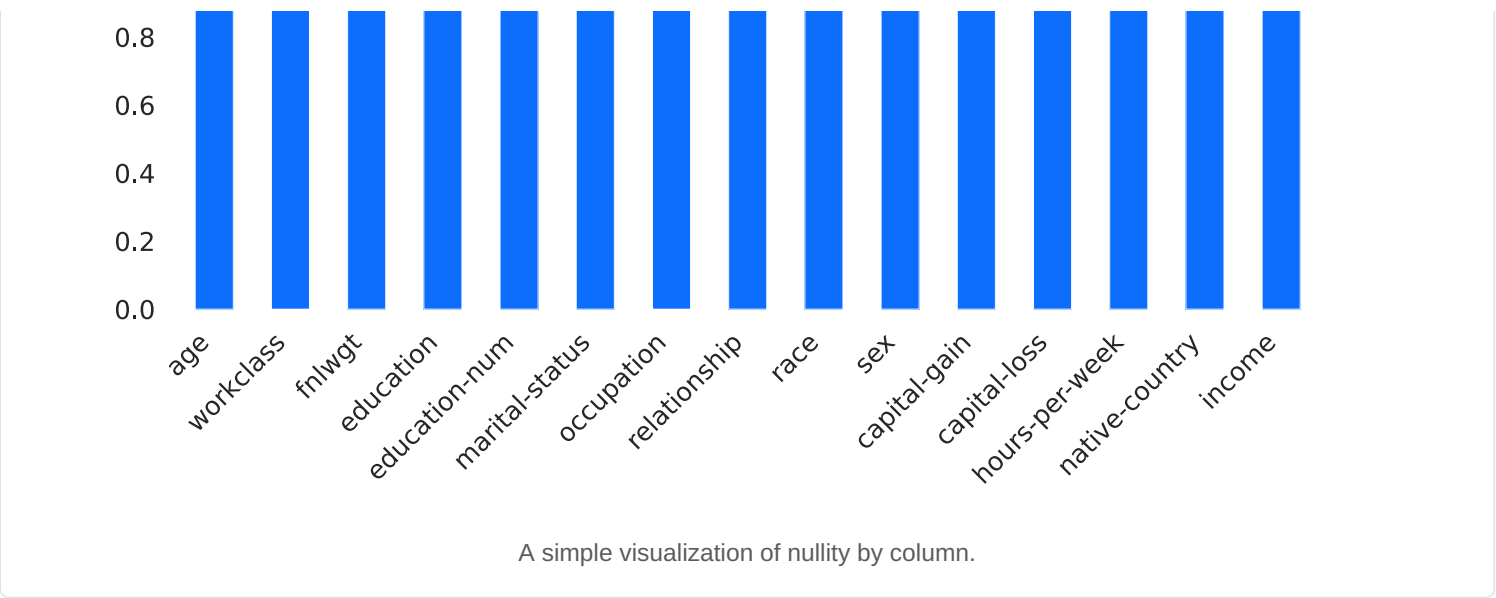


Correlations



Missing values





Sample

First rows

Last rows

	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	
0	39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	\
1	50	Self-emp-not-inc	83311	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	\
2	38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	\
3	53	Private	234721	11th	7	Married-civ-spouse	Handlers-cleaners	Husband	E
4	28	Private	338409	Bachelors	13	Married-civ-spouse	Prof-specialty	Wife	E
5	37	Private	284582	Masters	14	Married-civ-spouse	Exec-managerial	Wife	\
6	49	Private	160187	9th	5	Married-spouse-absent	Other-service	Not-in-family	E
7	52	Self-emp-not-inc	209642	HS-grad	9	Married-civ-spouse	Exec-managerial	Husband	\
8	31	Private	45781	Masters	14	Never-married	Prof-specialty	Not-in-family	\
9	42	Private	159449	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	\

Duplicate rows

Most frequently occurring

	age	workclass	fnlwgt	education	education-num	marital-status	occupation	relationship	race	sex	capital-gain	cap
10	25	Private	195994	1st-4th	2	Never-married	Priv-house-serv	Not-in-family	White	Female	0	0
0	18	Self-emp-inc	378036	12th	8	Never-married	Farming-fishing	Own-child	White	Male	0	0
1	19	Private	97261	HS-grad	9	Never-married	Farming-fishing	Not-in-family	White	Male	0	0

2	19	Private	138153	Some-college	10	Never-married	Adm-clerical	Own-child	White	Female	0	0
3	19	Private	146679	Some-college	10	Never-married	Exec-managerial	Own-child	Black	Male	0	0
4	19	Private	251579	Some-college	10	Never-married	Other-service	Own-child	White	Male	0	0
5	20	Private	107658	Some-college	10	Never-married	Tech-support	Not-in-family	White	Female	0	0
6	21	Private	243368	Preschool	1	Never-married	Farming-fishing	Not-in-family	White	Male	0	0
7	21	Private	250051	Some-college	10	Never-married	Prof-specialty	Own-child	White	Female	0	0
8	23	Private	240137	5th-6th	3	Never-married	Handlers-cleaners	Not-in-family	White	Male	0	0

