

Overview

Dataset Statistics

Number of Variables	15
Number of Rows	32561
Missing Cells	0
Missing Cells (%)	0.0%
Duplicate Rows	24
Duplicate Rows (%)	0.1%
Total Size in Memory	19.9 MB
Average Row Size in Memory	640.3 B
Variable Types	Numerical: 6 Categorical: 9

Dataset Insights

fnlwgt is skewed	Skewed
education_num is skewed	Skewed
capital_gain is skewed	Skewed
capital_loss is skewed	Skewed
hours_per_week is skewed	Skewed
capital_gain has 29849 (91.67%) zeros	Zeros
capital_loss has 31042 (95.33%) zeros	Zeros

Variables

Sort by

Feature order

☐ Reverse order

age

numerical

Show Details

Approximate Distinct Count	73	Mean	38.5816
Approximate Unique (%)	0.2%	Minimum	17
Missing	0	Maximum	90
Missing (%)	0.0%	Zeros	0
Infinite	0	Zeros (%)	0.0%
Infinite (%)	0.0%	Negatives	0
Memory Size	520976	Negatives (%)	0.0%

age

workclass

categorical

Show Details

Approximate Distinct Count	9
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	2372538

workclass

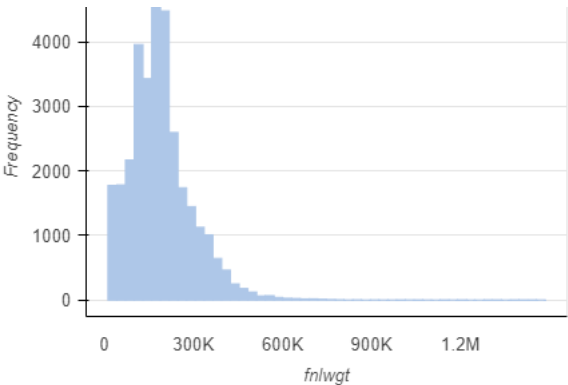
fnlwgt

fnlwgt
numerical

Show Details

Approximate Distinct Count	21648
Approximate Unique (%)	66.5%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Memory Size	520976

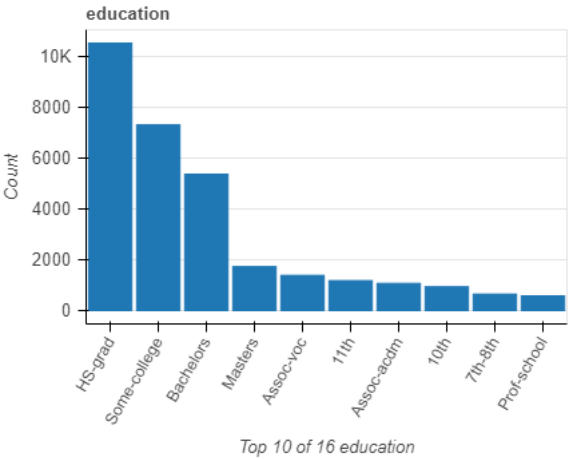
Mean	189778.3665
Minimum	12285
Maximum	1484705
Zeros	0
Zeros (%)	0.0%
Negatives	0
Negatives (%)	0.0%



education
categorical

Show Details

Approximate Distinct Count	16
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	2391075

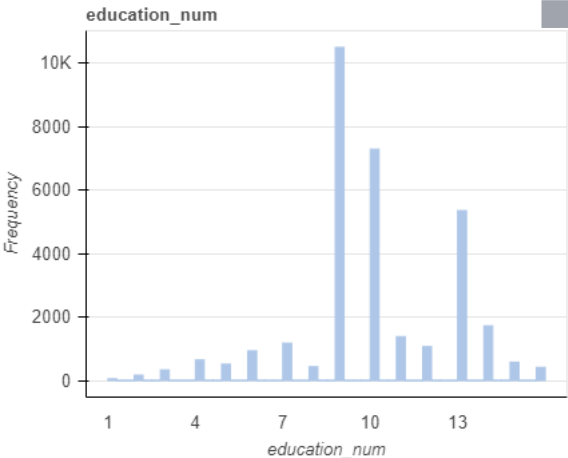


education_num
numerical

Show Details

Approximate Distinct Count	16
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Memory Size	520976

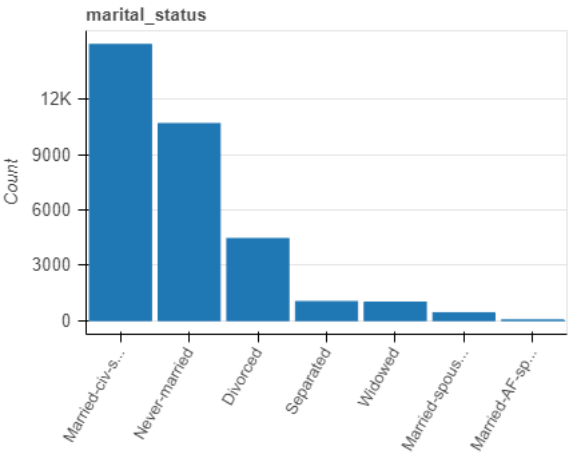
Mean	10.0807
Minimum	1
Maximum	16
Zeros	0
Zeros (%)	0.0%
Negatives	0
Negatives (%)	0.0%



marital_status
categorical

Show Details

Approximate Distinct Count	7
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	2585801



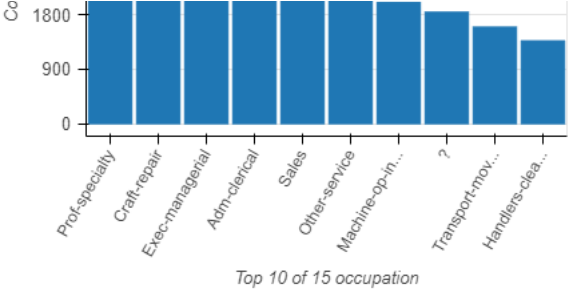
Approximate Distinct Count	15
----------------------------	----



occupation
categorical

Show Details

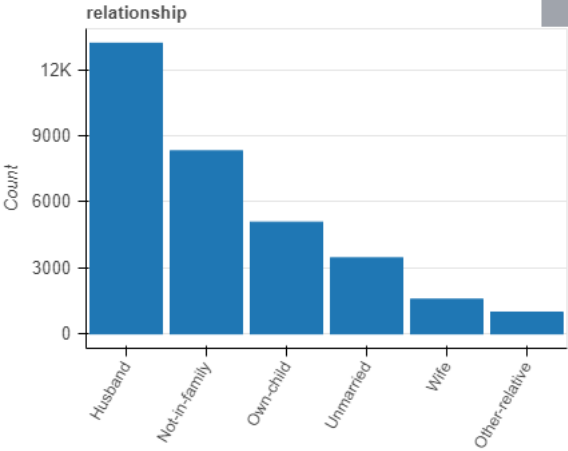
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	2513771



relationship
categorical

Show Details

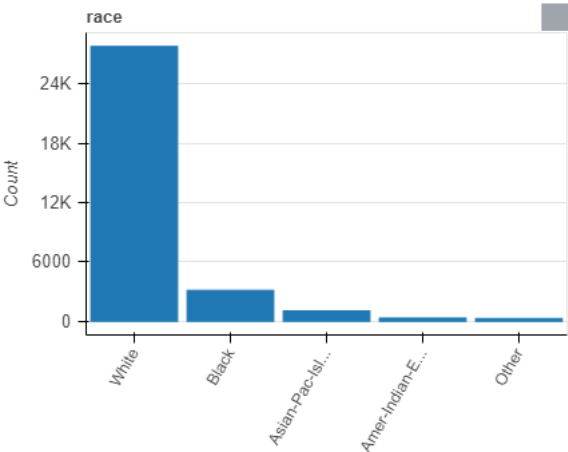
Approximate Distinct Count	6
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	2413413



race
categorical

Show Details

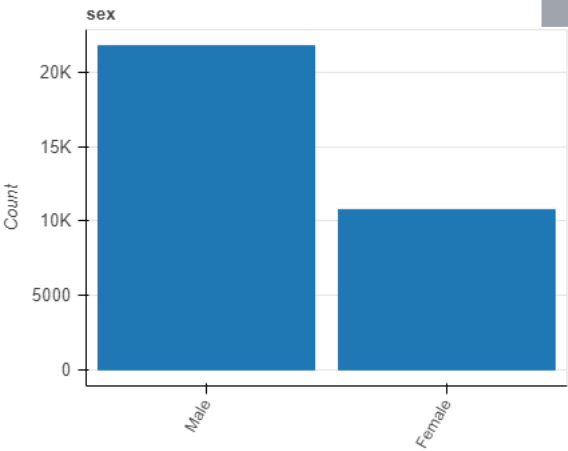
Approximate Distinct Count	5
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	2296820



sex
categorical

Show Details

Approximate Distinct Count	2
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	2268251



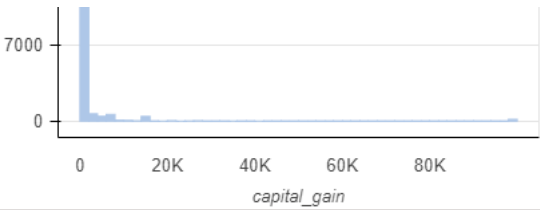
capital_gain
numerical

Approximate Distinct Count	119	Infinite	0
Approximate Unique (%)	0.4%	Infinite (%)	0.0%
Missing	0	Memory Size	520976
Missing (%)	0.0%	Mean	1077.6488



Show Details

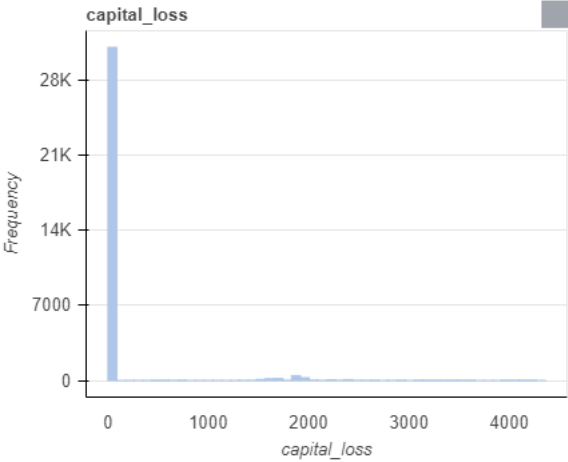
Minimum	0	Zeros (%)	91.7%
Maximum	99999	Negatives	0
Zeros	29849	Negatives (%)	0.0%



capital_loss
numerical

Show Details

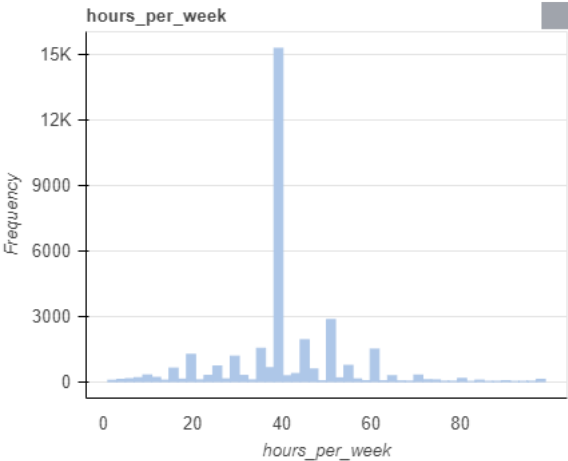
Approximate Distinct Count	92	Mean	87.3038
Approximate Unique (%)	0.3%	Minimum	0
Missing	0	Maximum	4356
Missing (%)	0.0%	Zeros	31042
Infinite	0	Zeros (%)	95.3%
Infinite (%)	0.0%	Negatives	0
Memory Size	520976	Negatives (%)	0.0%



hours_per_week
numerical

Show Details

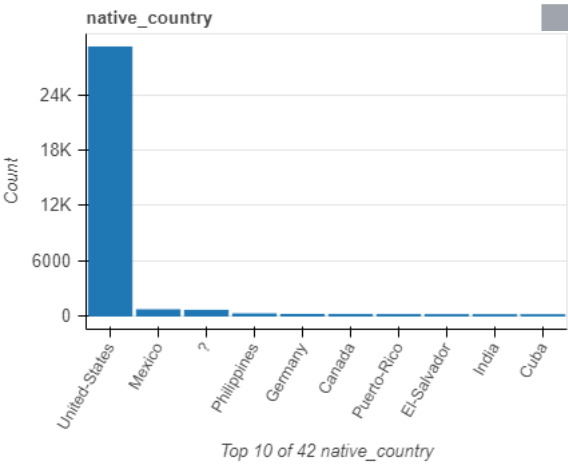
Approximate Distinct Count	94	Mean	40.4375
Approximate Unique (%)	0.3%	Minimum	1
Missing	0	Maximum	99
Missing (%)	0.0%	Zeros	0
Infinite	0	Zeros (%)	0.0%
Infinite (%)	0.0%	Negatives	0
Memory Size	520976	Negatives (%)	0.0%



native_country
categorical

Show Details

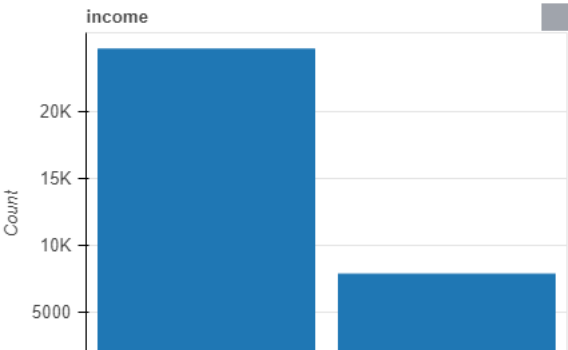
Approximate Distinct Count	42
Approximate Unique (%)	0.1%
Missing	0
Missing (%)	0.0%
Memory Size	2516765



income
categorical

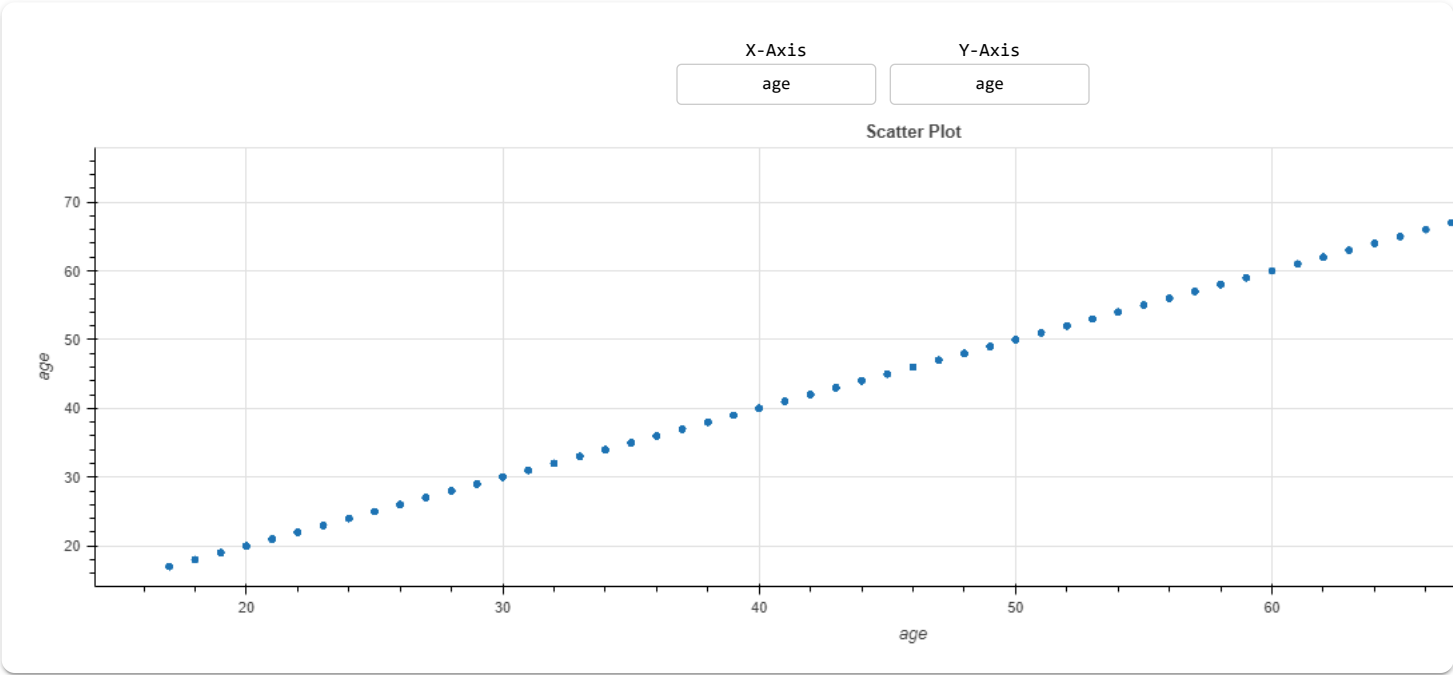
Show Details

Approximate Distinct Count	2
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	2271429

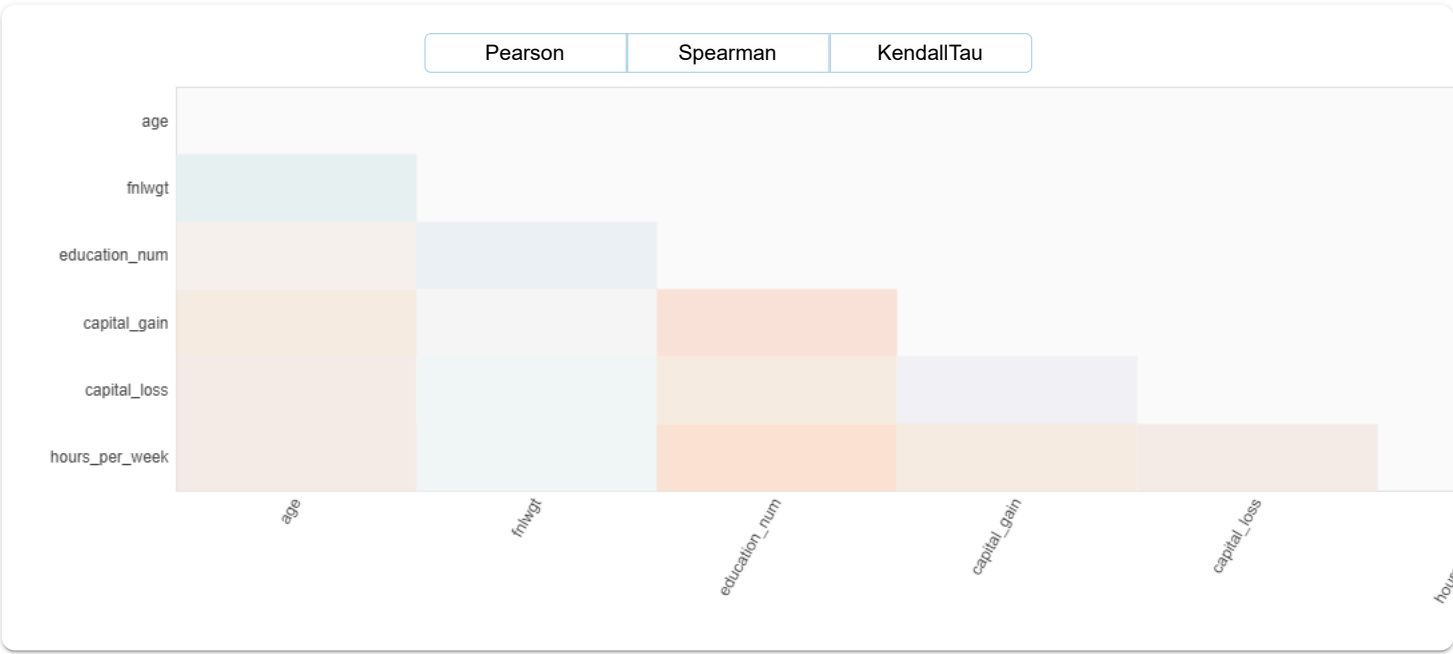




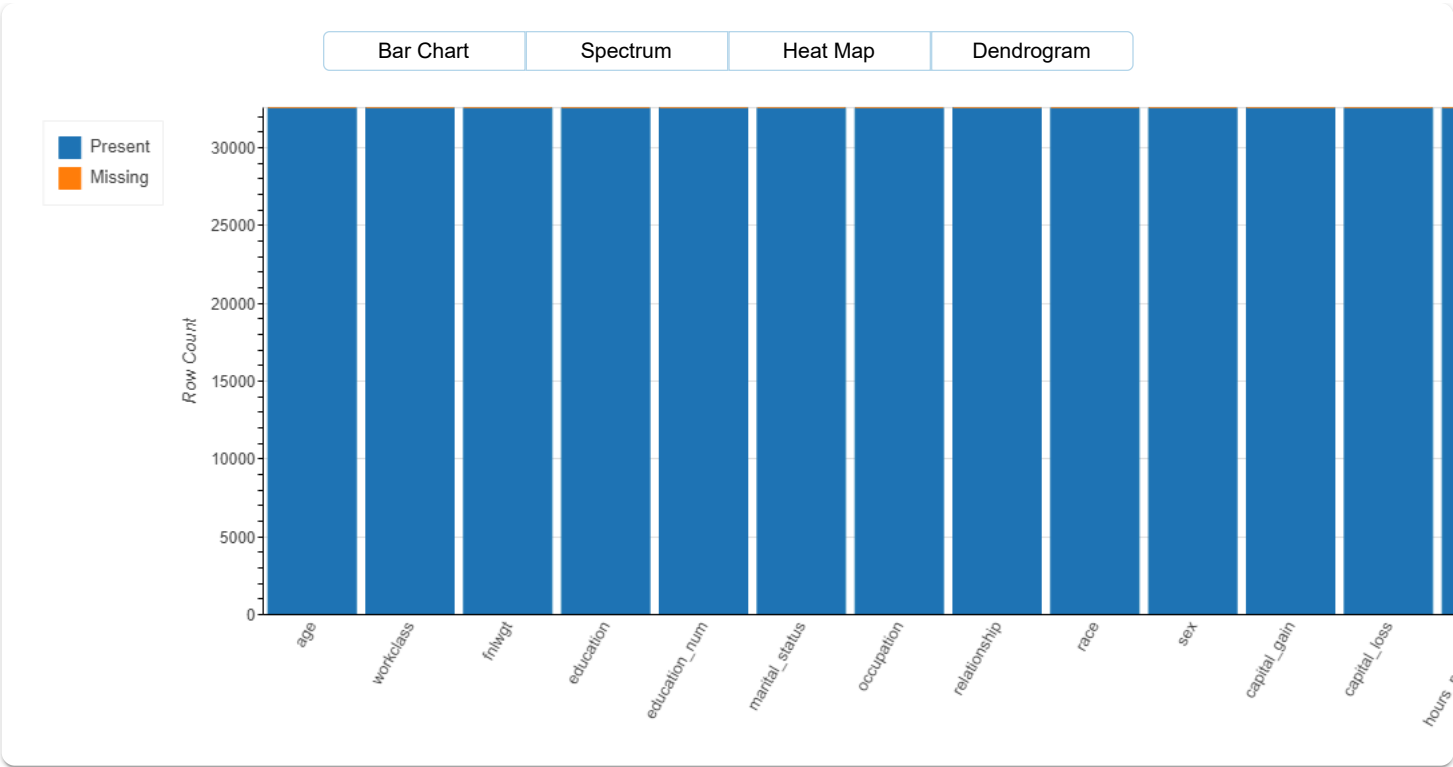
Interactions



Correlations



Missing Values



Report generated with [DataPrep](#)