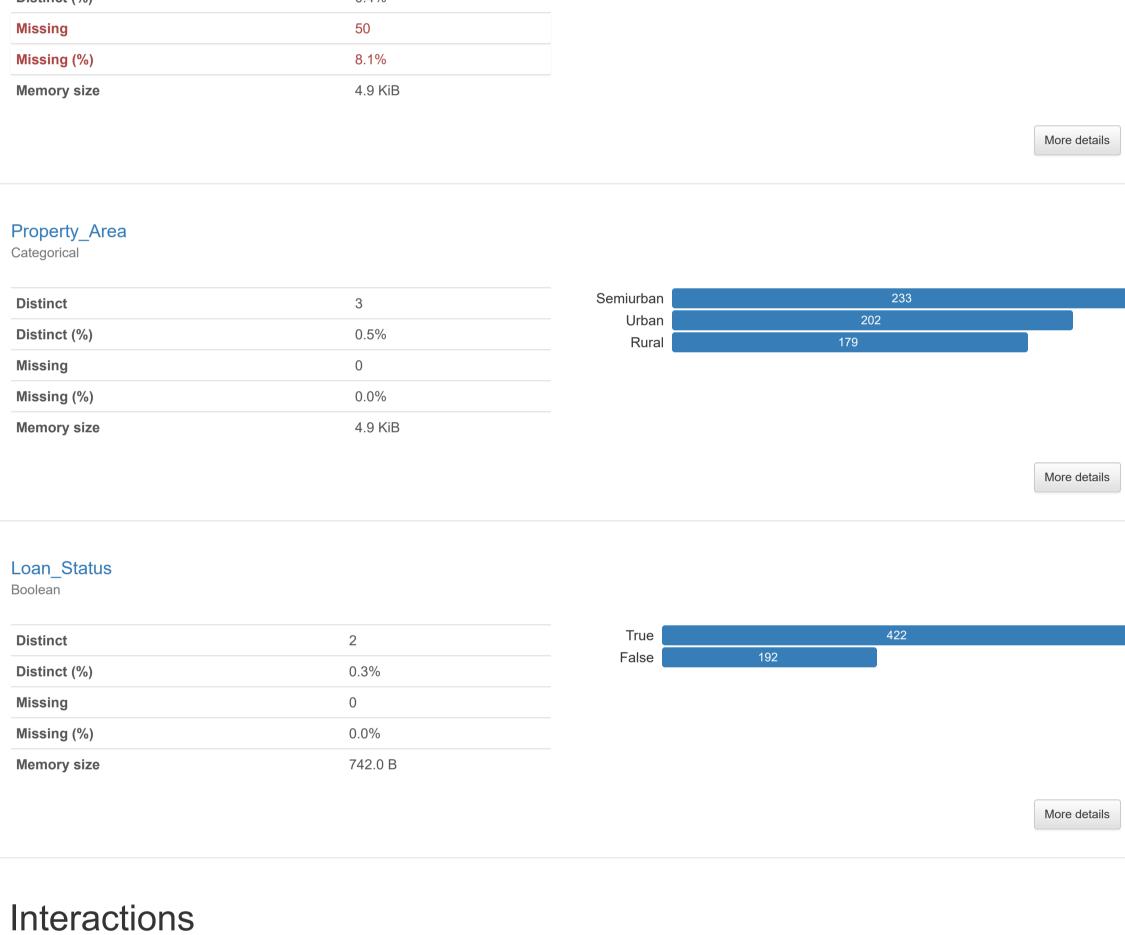
Pandas Profiling Report Overview Variables Interactions Correlations Missing values Sample Overview Alerts 14 Overview Reproduction Variable types Dataset statistics Number of variables Categorical 6 13 3 Number of observations 614 Boolean Missing cells 149 Numeric 4 1.9% Missing cells (%) 0 **Duplicate rows** Duplicate rows (%) 0.0% 62.5 KiB **Total size in memory** Average record size in memory 104.2 B Variables Select Columns Loan\_ID Categorical HIGH CARDINALITY UNIFORM UNIQUE LP001002 1 614 **Distinct** LP002328 Distinct (%) 100.0% LP002305 LP002308 0 Missing LP002314 0.0% Missing (%) Other values ... 609 4.9 KiB Memory size More details Gender Categorical Male Distinct 2 Female Distinct (%) 0.3% 13 Missing 2.1% Missing (%) Memory size 4.9 KiB More details Married Boolean 398 True 2 **Distinct** False Distinct (%) 0.3% (Missing) 3 3 Missing Missing (%) 0.5% **Memory size** 1.3 KiB More details Dependents Categorical 345 0 **Distinct** 4 Distinct (%) 0.7% 101 2 Missing 15 3+ 2.4% Missing (%) 4.9 KiB Memory size More details Education Categorical 480 Graduate 2 **Distinct** Not Graduate Distinct (%) 0.3% 0 Missing Missing (%) 0.0% 4.9 KiB **Memory size** More details Self\_Employed Boolean 500 False **Distinct** 2 True Distinct (%) 0.3% (Missing) 32 Missing 32 Missing (%) 5.2% Memory size 1.3 KiB More details ApplicantIncome Real number (R) 505 150 **Distinct** Minimum Distinct (%) 82.2% Maximum 81000 Missing 0 0 Zeros Missing (%) 0.0% 0.0% Zeros (%) 0 0 Infinite Negative Infinite (%) 0.0% Negative (%) 0.0% Mean 5403.4593 Memory size 4.9 KiB More details CoapplicantIncome Real number ( $\mathbb{R}$ ) 287 0 **Distinct** Minimum Distinct (%) 46.7% Maximum 41667 Missing 0 273 Zeros 0.0% 44.5% Missing (%) Zeros (%) 0 Infinite 0 Negative 0.0% Infinite (%) 0.0% Negative (%) Mean 1621.2458 4.9 KiB Memory size More details LoanAmount Real number  $(\mathbb{R})$ HIGH CORRELATION MISSING 203 9 **Distinct** Minimum Distinct (%) 34.3% Maximum 700 Missing 22 Zeros 0 Missing (%) 3.6% 0.0% Zeros (%) Infinite 0 0 Negative Infinite (%) 0.0% Negative (%) 0.0% Mean 146.41216 Memory size 4.9 KiB More details Loan\_Amount\_Term Real number  $(\mathbb{R})$ Distinct 10 Minimum 12 480 1.7% Maximum Distinct (%) 14 0 Missing Zeros Missing (%) 2.3% 0.0% Zeros (%) Infinite 0 Negative 0 Infinite (%) 0.0% Negative (%) 0.0% Mean 342 Memory size 4.9 KiB More details Credit\_History Categorical HIGH CORRELATION MISSING 1.0 2 **Distinct** 0.0 Distinct (%) 0.4% Missing 50 Missing (%) 8.1% 4.9 KiB **Memory size** 



## 400

ApplicantIncome

Loan\_Amount\_Term

500

CoapplicantIncome

ApplicantIncome

LoanAmount

Gender

Loan\_Amount\_Term

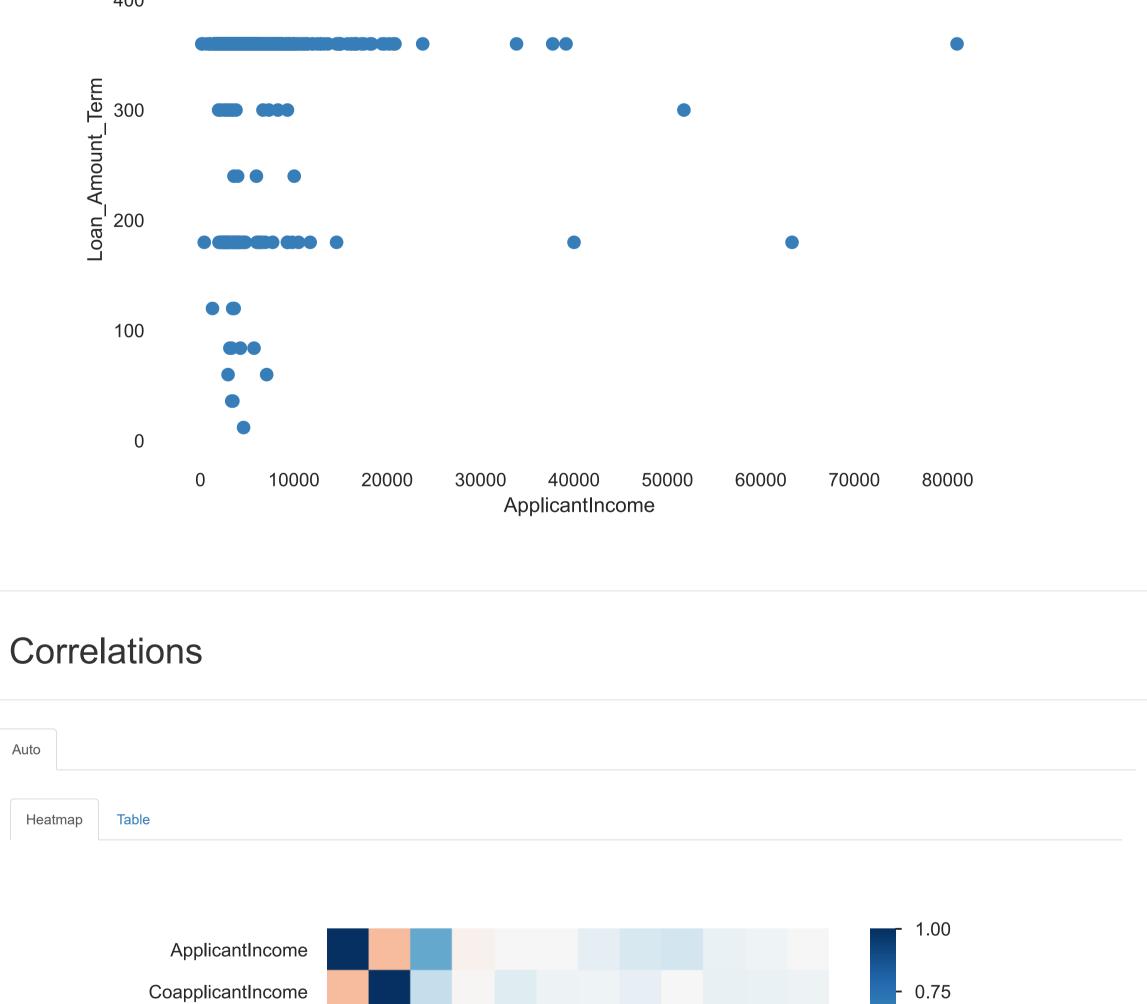
Sample

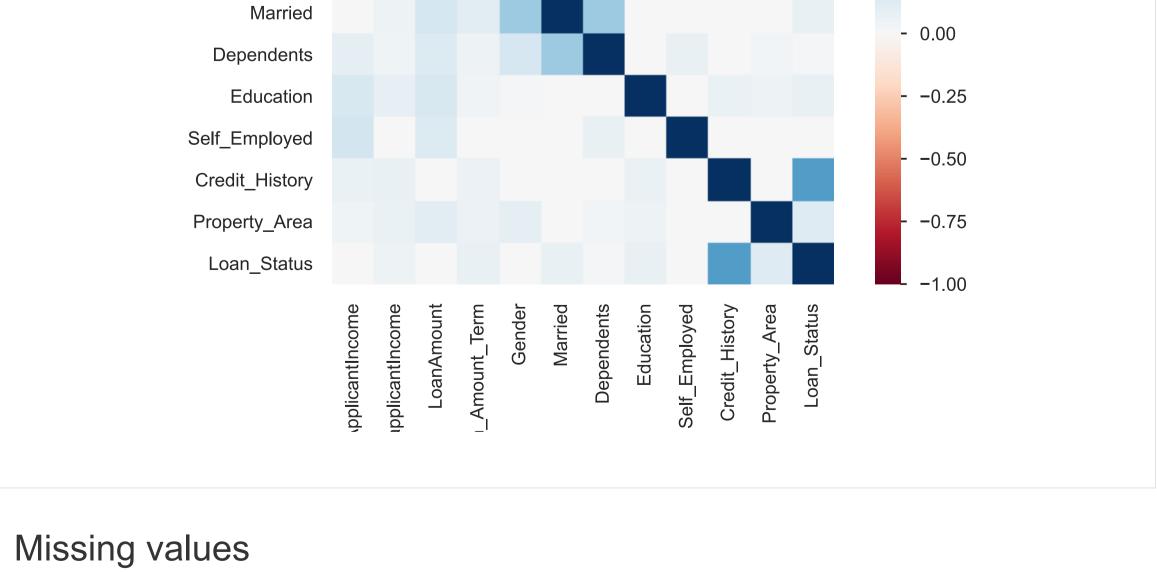
LoanAmount

CoapplicantIncome

Loan\_Amount\_Term

LoanAmount





0.50

0.25

## Count Matrix Heatmap 1.0 0.8 0.6 0.4 0.2 0.0 1.con ID contest Marked Department Lettershift Lettershift Contest Marked Lettershift Contest Marked Lettershift Lettershift Contest Marked Lettershift Lettershift Lettershift Contest Marked Lettershift Lettershi

	Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Terr
0	LP001002	Male	No	0	Graduate	No	5849	0.0	NaN	360.0
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128.0	360.0
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66.0	360.0
3	LP001006	Male	Yes	0	Not Graduate	No	2583	2358.0	120.0	360.0
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141.0	360.0
5	LP001011	Male	Yes	2	Graduate	Yes	5417	4196.0	267.0	360.0
6	LP001013	Male	Yes	0	Not Graduate	No	2333	1516.0	95.0	360.0
7	LP001014	Male	Yes	3+	Graduate	No	3036	2504.0	158.0	360.0
8	LP001018	Male	Yes	2	Graduate	No	4006	1526.0	168.0	360.0
9	LP001020	Male	Yes	1	Graduate	No	12841	10968.0	349.0	360.0