

## Input Variables:

- Loan\_ID
- Customer
- Date
- Gender
- Married
- Dependents
- Age
- Application\_Type
- Education
- Self\_Employed
- ApplicantIncome
- CoapplicantIncome
- LoanAmount
- Loan\_Amount\_Term
- Credit\_History
- Property\_Area
- Application\_Number
- Amount\_Request

## Columns Removed:

- **Amount\_Request** – This column is a duplicate of the Loan Amount Column and there is no need to have the same data represented in more than one column.
- **Loan\_ID** – Cardinality is the same as count so no information of value can be extracted from this data.
- **Customer** – Cardinality is the same as count so no information of value can be extracted from this data.
- **Application\_Number** – Cardinality is the same as count so no information of value can be extracted from this data.
- **Application\_Type** – Cardinality is one count so no information of value can be extracted from this data.

## Categorical Data:

- Loan\_ID
- Customer
- Date
- Gender
- Married
- Application\_Type
- Education
- Self\_Employed
- Credit\_History
- Property\_Area
- Application\_Number

	A	B	C	D	E	F	G	H	I
1	Features	Distinct Count	Missing%	Cardinality	Mode	Mode Freq.	DataType	DataSource	TotalRows
2	Loan_ID	614	0.00000	614	N/A	N/A	Categorical	Default Data Source	614
3	Customer	614	0.00000	614	N/A	N/A	Categorical	Default Data Source	614
4	Date	614	0.00000	576	8/16/2023	9	Categorical	Default Data Source	614
5	Gender	601	0.02117	2	Male	489	Categorical	Default Data Source	614
6	Married	611	0.00489	2	Yes	398	Categorical	Default Data Source	614
7	Application_Type	614	0.00000	1	Loan	614	Categorical	Default Data Source	614
8	Education	614	0.00000	2	Graduate	480	Categorical	Default Data Source	614
9	Self_Employed	582	0.05212	2	No	500	Categorical	Default Data Source	614
10	Credit_History	564	0.08143	2	1	475	Categorical	Default Data Source	614
11	Property_Area	614	0.00000	3	Semiurban	233	Categorical	Default Data Source	614
12	Application_Number	614	0.00000	614	N/A	N/A	Categorical	Default Data Source	614

## Continuous Data:

- Dependents
- Age
- ApplicantIncome
- CoapplicantIncome
- LoanAmount
- Loan\_Amount\_Term
- Amount\_Request

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Features	Count	Missing%	Cardinality	Min	Max	Median	Mode	Mean	Std Dev.	DataType	DataSource	TotalRows
2	Dependents	599	0.02443	15	0	100	0	0	1.11519	4.56998	Continuous	Default Data Source	614
3	Age	513	0.1645	21	-2	180	41	55	37.35673	15.10521	Continuous	Default Data Source	614
4	ApplicantIncome	614	0	505	150	81000	3812.5	2500	5403.45928	6109.0417	Continuous	Default Data Source	614
5	CoapplicantIncome	614	0	287	0	41667	1188.5	0	1621.2458	2926.2483	Continuous	Default Data Source	614
6	LoanAmount	592	0.03583	203	9	700	128	120	146.41216	85.5873	Continuous	Default Data Source	614
7	Loan_Amount_Term	600	0.0228	10	12	480	360	360	342	65.1204	Continuous	Default Data Source	614
8	Amount_Requested	592	0.03583	203	9	700	128	120	146.41216	85.5873	Continuous	Default Data Source	614

