

The plot will not show in a notebook environment, please try 'show\_browser' if you want to open it in browser

DataPrep Report

OverviewVariables ≡InteractionsCorrelationsMissing Values

Overview

Dataset Statistics

Number of Variables	28
Number of Rows	100000
Missing Cells	60071
Missing Cells (%)	2.1%
Duplicate Rows	0
Duplicate Rows (%)	0.0%
Total Size in Memory	135.5 MB
Average Row Size in Memory	1.4 KB
Variable Types	Categorical: 20 Numerical: 8

Dataset Insights

Name	has 9985 (9.98%) missing values	Missing
Monthly_Inhand_Salary	has 15002 (15.0%) missing values	Missing
Type_of_Loan	has 11408 (11.41%) missing values	Missing
Num_of_Delayed_Payment	has 7002 (7.0%) missing values	Missing
Num_Credit_Inquiries	has 1965 (1.96%) missing values	Missing
Credit_History_Age	has 9030 (9.03%) missing values	Missing
Amount_invested_monthly	has 4479 (4.48%) missing values	Missing
Monthly_Balance	has 1200 (1.2%) missing values	Missing
Num_Bank_Accounts	is skewed	Skewed
Num_Credit_Card	is skewed	Skewed

1

2

3

Variables

Sort by

Feature order

Reverse order

ID

categorical

Hide Details

Approximate Distinct Count	100000
Approximate Unique (%)	100.0%
Missing	0
Missing (%)	0.0%
Memory Size	6.8 MB

Stats

Pie Chart

Word Frequency

Word Length

Length

Mean	6.6006
Standard Deviation	0.4898
Median	7
Minimum	6
Maximum	7

Sample

1st row	0x1602
2nd row	0x1603
3rd row	0x1604
4th row	0x1605
5th row	0x1606

Letter

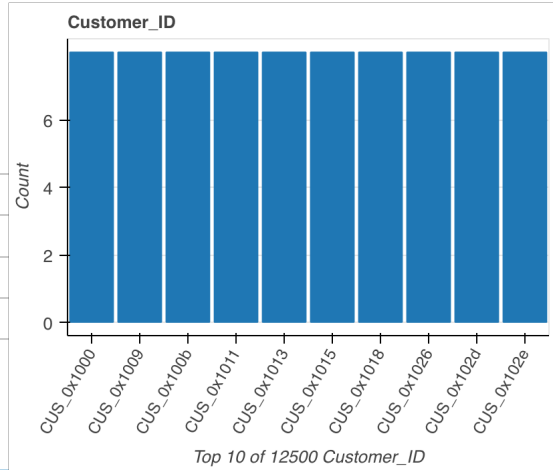
Count	245640
Lowercase Letter	245640
Space Separator	0
Uppercase Letter	0
Dash Punctuation	0
Decimal Number	414424

ID

Top 10 of 100000 ID

**Customer\_ID**  
categorical[Hide Details](#)

Approximate Distinct Count	12500
Approximate Unique (%)	12.5%
Missing	0
Missing (%)	0.0%
Memory Size	7.1 MB



Stats

Pie Chart

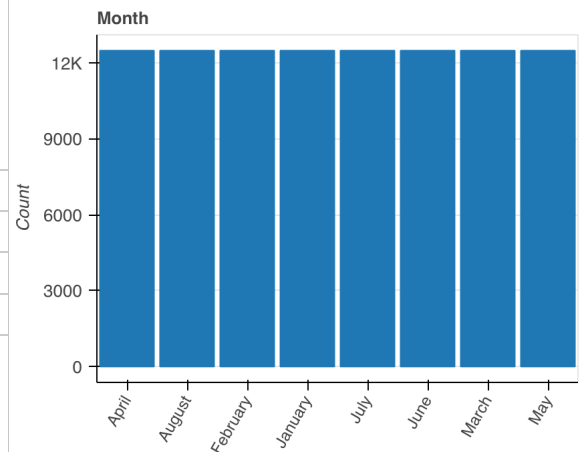
Word Frequency

Word Length

Length		Sample		Letter	
Mean	9.9395	1st row	CUS_0xd40	Count	532744
Standard Deviation	0.2384	2nd row	CUS_0xd40	Lowercase Letter	232744
Median	10	3rd row	CUS_0xd40	Space Separator	0
Minimum	9	4th row	CUS_0xd40	Uppercase Letter	300000
Maximum	10	5th row	CUS_0xd40	Dash Punctuation	0
				Decimal Number	361208

**Month**  
categorical[Hide Details](#)

Approximate Distinct Count	8
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6.7 MB



Stats

Pie Chart

Word Frequency

Word Length

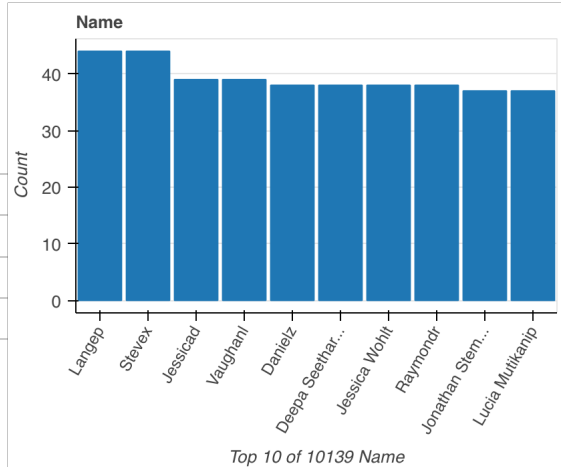
Length		Sample		Letter	
Mean	5.25	1st row	January	Count	525000
Standard Deviation	1.5613	2nd row	February	Lowercase Letter	425000
Median	5	3rd row	March	Space Separator	0
Minimum	3	4th row	April	Uppercase Letter	100000
Maximum	8	5th row	May	Dash Punctuation	0
				Decimal Number	0

**Name**

categorical

Hide Details

Approximate Distinct Count	10139
Approximate Unique (%)	11.3%
Missing	9985
Missing (%)	10.0%
Memory Size	6.4 MB



Stats

Pie Chart

Word Frequency

Word Length

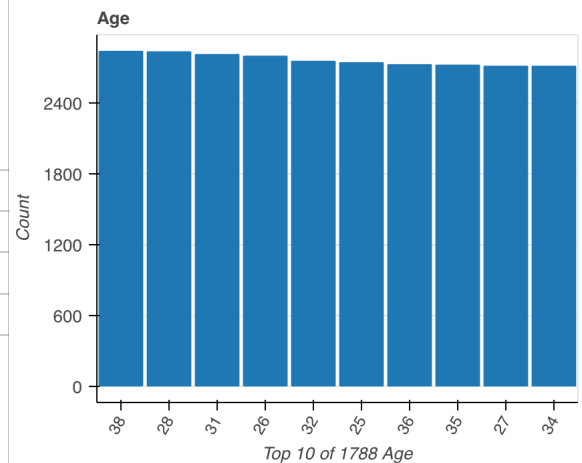
Length		Sample		Letter	
Mean	9.7698	1st row	Aaron Maashoh	Count	840131
Standard Deviation	4.2389	2nd row	Aaron Maashoh	Lowercase Letter	714761
Median	9	3rd row	Aaron Maashoh	Space Separator	35698
Minimum	2	4th row	Aaron Maashoh	Uppercase Letter	125370
Maximum	25	5th row	Aaron Maashoh	Dash Punctuation	1457
				Decimal Number	0

**Age**

categorical

Hide Details

Approximate Distinct Count	1788
Approximate Unique (%)	1.8%
Missing	0
Missing (%)	0.0%
Memory Size	6.4 MB



Stats

Pie Chart

Word Frequency

Word Length

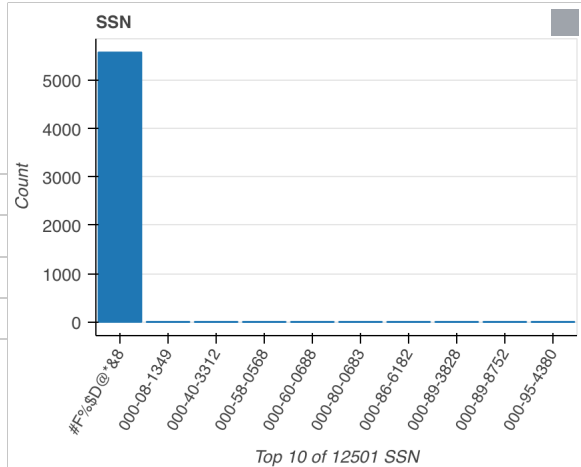
Length		Sample		Letter	
Mean	2.1029	1st row	23	Count	0
Standard Deviation	0.3835	2nd row	23	Lowercase Letter	0
Median	2	3rd row	-500	Space Separator	0
Minimum	2	4th row	23	Uppercase Letter	0
Maximum	5	5th row	23	Dash Punctuation	886
				Decimal Number	204465

## SSN

categorical

[Hide Details](#)

Approximate Distinct Count	12501
Approximate Unique (%)	12.5%
Missing	0
Missing (%)	0.0%
Memory Size	7.2 MB



Stats	Pie Chart	Word Frequency	Word Length
-------	-----------	----------------	-------------

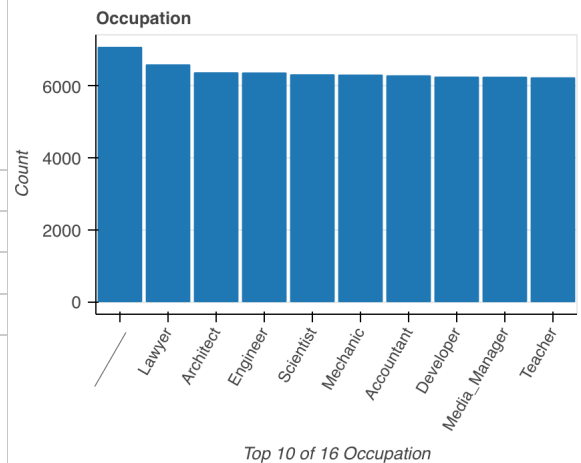
Length		Sample		Letter	
Mean	10.8886	1st row	821-00-0265	Count	11144
Standard Deviation	0.4588	2nd row	821-00-0265	Lowercase Letter	0
Median	11	3rd row	821-00-0265	Space Separator	0
Minimum	9	4th row	821-00-0265	Uppercase Letter	11144
Maximum	11	5th row	821-00-0265	Dash Punctuation	188856
				Decimal Number	855424

## Occupation

categorical

[Hide Details](#)

Approximate Distinct Count	16
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	7.0 MB

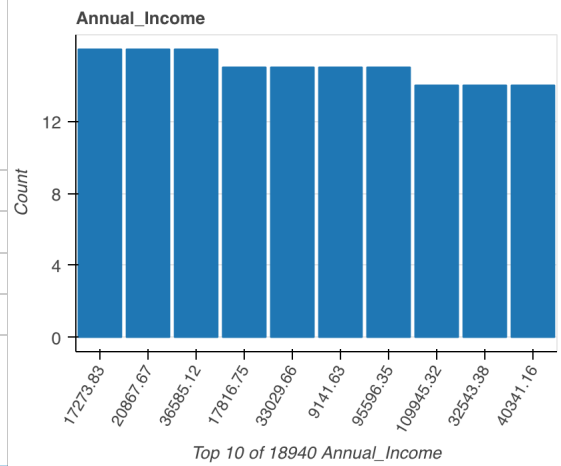


Stats	Pie Chart	Word Frequency	Word Length
-------	-----------	----------------	-------------

Length		Sample		Letter	
Mean	8.4311	1st row	Scientist	Count	787447
Standard Deviation	1.9944	2nd row	Scientist	Lowercase Letter	688277
Median	8	3rd row	Scientist	Space Separator	0
Minimum	6	4th row	Scientist	Uppercase Letter	99170
Maximum	13	5th row	Scientist	Dash Punctuation	0
				Decimal Number	0

**Annual\_Income**  
categorical[Hide Details](#)

Approximate Distinct Count	18940
Approximate Unique (%)	18.9%
Missing	0
Missing (%)	0.0%
Memory Size	7.0 MB



Stats

Pie Chart

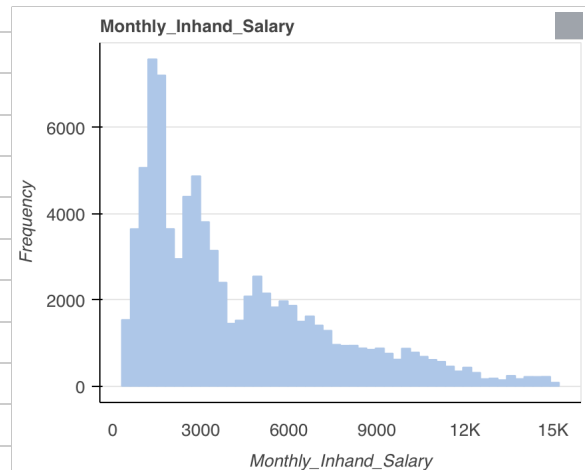
Word Frequency

Word Length

Length		Sample		Letter	
Mean	8.3091	1st row	19114.12	Count	0
Standard Deviation	1.4473	2nd row	19114.12	Lowercase Letter	0
Median	8	3rd row	19114.12	Space Separator	0
Minimum	6	4th row	19114.12	Uppercase Letter	0
Maximum	19	5th row	19114.12	Dash Punctuation	0
				Decimal Number	723934

**Monthly\_Inhand\_S...**  
numerical[Hide Details](#)

Approximate Distinct Count	13235
Approximate Unique (%)	15.6%
Missing	15002
Missing (%)	15.0%
Infinite	0
Infinite (%)	0.0%
Memory Size	1.3 MB
Mean	4194.1708
Minimum	303.6454
Maximum	15204.6333
Zeros	0
Zeros (%)	0.0%
Negatives	0
Negatives (%)	0.0%

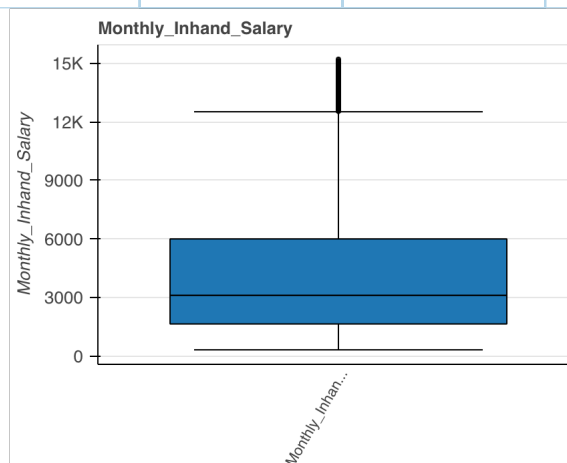


Stats

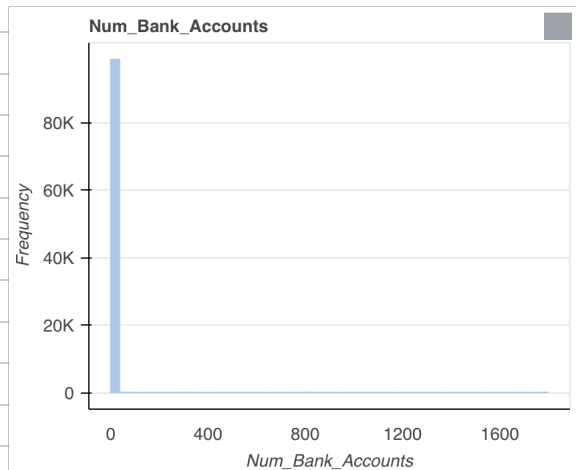
KDE Plot

Normal Q-Q Plot

Box Plot



<b>Num_Bank_Accounts</b> numerical	Approximate Distinct Count	943
	Approximate Unique (%)	0.9%
	Missing	0
	Missing (%)	0.0%
	Infinite	0
	Infinite (%)	0.0%
	Memory Size	1.5 MB
	Mean	17.0913
	Minimum	-1
	Maximum	1798
	Zeros	4328
	Zeros (%)	4.3%
	Negatives	21

[Hide Details](#)

Stats

KDE Plot

Normal Q-Q Plot

Box Plot

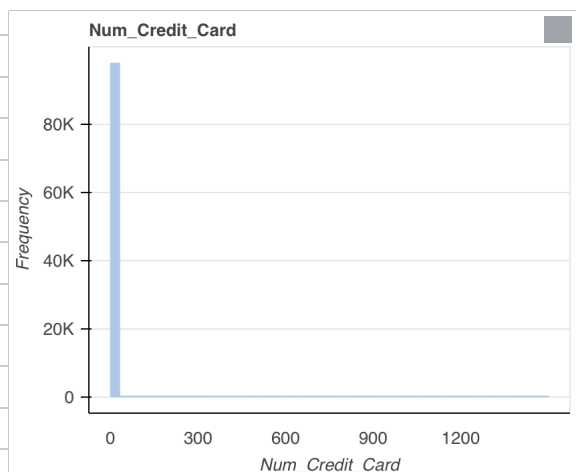
## Quantile Statistics

## Descriptive Statistics

Minimum	-1
5-th Percentile	1
Q1	3
Median	6
Q3	7
95-th Percentile	10
Maximum	1798
Range	1799
IQR	4

Mean	17.0913
Standard Deviation	117.4048
Variance	13783.8951
Sum	1.7091×10 <sup>06</sup>
Skewness	11.2023
Kurtosis	132.5033
Coefficient of Variation	6.8693

<b>Num_Credit_Card</b> numerical	Approximate Distinct Count	1179
	Approximate Unique (%)	1.2%
	Missing	0
	Missing (%)	0.0%
	Infinite	0
	Infinite (%)	0.0%
	Memory Size	1.5 MB
	Mean	22.4744
	Minimum	0
	Maximum	1499
	Zeros	13
	Zeros (%)	0.0%
	Negatives	0

[Hide Details](#)

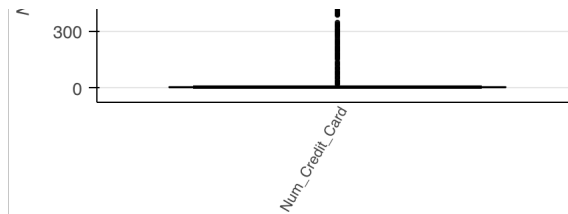
Stats

KDE Plot

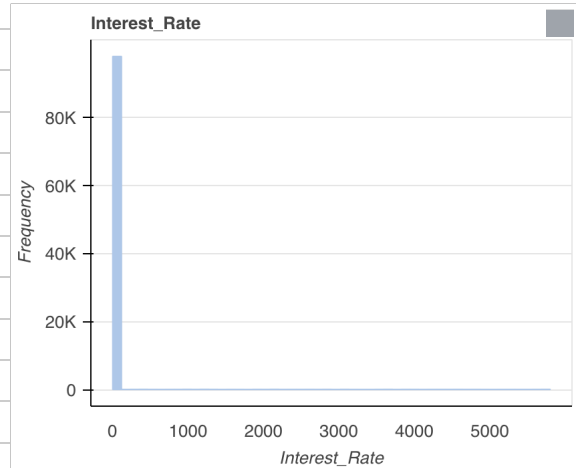
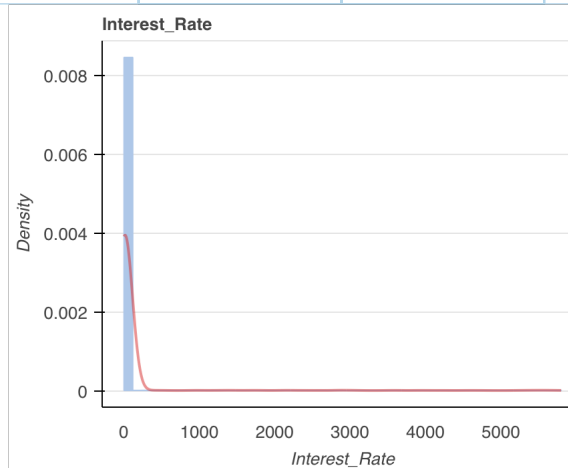
Normal Q-Q Plot

Box Plot

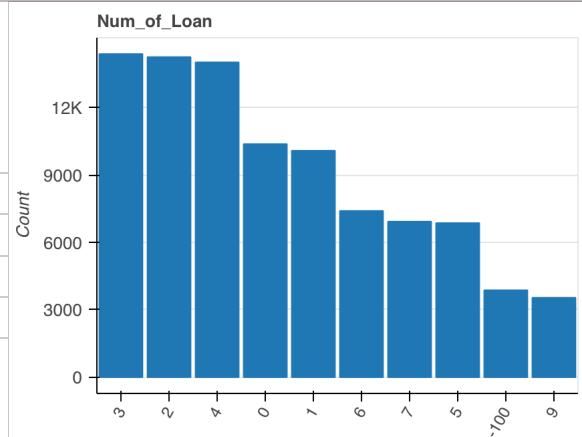


**Interest\_Rate**  
numerical[Hide Details](#)

Approximate Distinct Count	1750
Approximate Unique (%)	1.8%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Memory Size	1.5 MB
Mean	72.466
Minimum	1
Maximum	5797
Zeros	0
Zeros (%)	0.0%
Negatives	0
Negatives (%)	0.0%

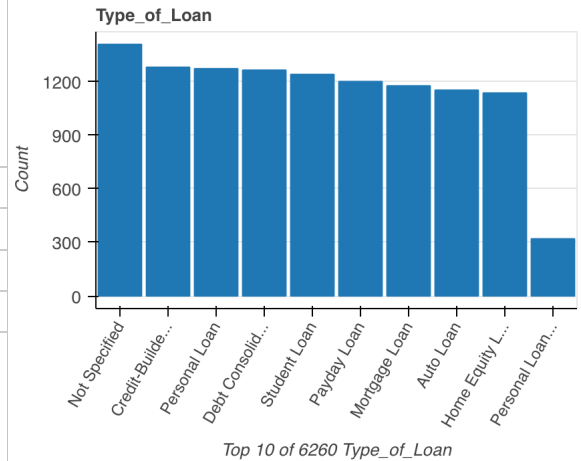
[Stats](#) [KDE Plot](#) [Normal Q-Q Plot](#) [Box Plot](#)**Num\_of\_Loan**  
categorical[Hide Details](#)

Approximate Distinct Count	434
Approximate Unique (%)	0.4%
Missing	0
Missing (%)	0.0%
Memory Size	6.3 MB

[Stats](#) [Pie Chart](#) [Word Frequency](#) [Word Length](#)

Length		Sample		Letter	
Mean	1.1748	1st row	4	Count	0
Standard Deviation	0.627	2nd row	4	Lowercase Letter	0
Median	1	3rd row	4	Space Separator	0
Minimum	1	4th row	4	Uppercase Letter	0
Maximum	5	5th row	4	Dash Punctuation	3876
				Decimal Number	108821

<b>Type_of_Loan</b> categorical <a href="#">Hide Details</a>	Approximate Distinct Count	6260
	Approximate Unique (%)	7.1%
	Missing	11408
	Missing (%)	11.4%
	Memory Size	11.1 MB



Stats

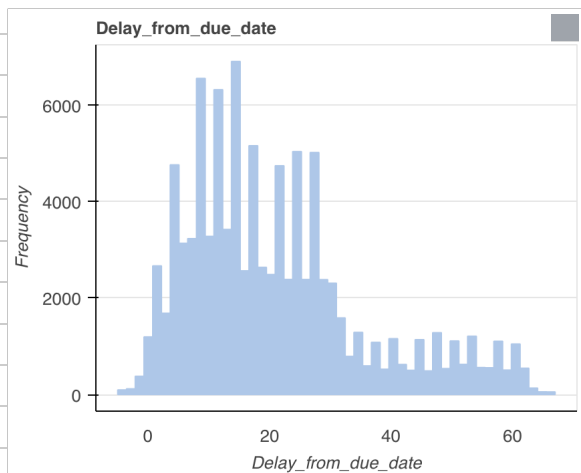
Pie Chart

Word Frequency

Word Length

Length		Sample		Letter	
Mean	66.6836	1st row	Auto Loan, Credit-...	Count	4829168
Standard Deviation	37.8881	2nd row	Auto Loan, Credit-...	Lowercase Letter	4004272
Median	60	3rd row	Auto Loan, Credit-...	Space Separator	773328
Minimum	9	4th row	Auto Loan, Credit-...	Uppercase Letter	824896
Maximum	182	5th row	Auto Loan, Credit-...	Dash Punctuation	40440
				Decimal Number	0

<b>Delay_from_due_d...</b> numerical <a href="#">Hide Details</a>	Approximate Distinct Count	73
	Approximate Unique (%)	0.1%
	Missing	0
	Missing (%)	0.0%
	Infinite	0
	Infinite (%)	0.0%
	Memory Size	1.5 MB
	Mean	21.0688
	Minimum	-5
	Maximum	67
	Zeros	1195
	Zeros (%)	1.2%
	Negatives	591
	Negatives (%)	0.6%



Stats

KDE Plot

Normal Q-Q Plot

Box Plot



## Quantile Statistics

Minimum	-5
5-th Percentile	3
Q1	10
Median	18
Q3	28
95-th Percentile	54
Maximum	67
Range	72
IQR	18

## Descriptive Statistics

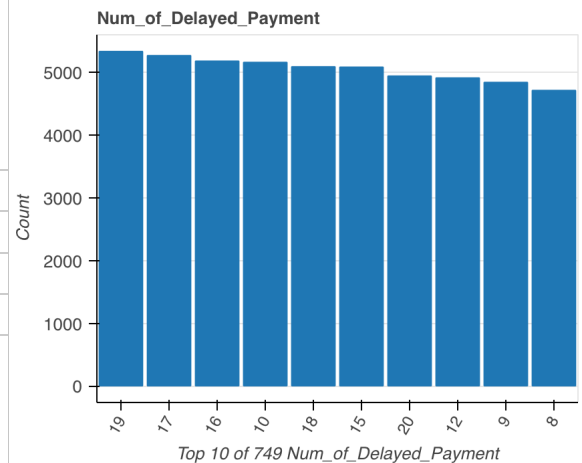
Mean	21.0688
Standard Deviation	14.8601
Variance	220.8227
Sum	2.1069×10 <sup>06</sup>
Skewness	0.9664
Kurtosis	0.3482
Coefficient of Variation	0.7053

Num\_of\_Delayed\_P...

categorical

Hide Details

Approximate Distinct Count	749
Approximate Unique (%)	0.8%
Missing	7002
Missing (%)	7.0%
Memory Size	5.9 MB



Stats

Pie Chart

Word Frequency

Word Length

## Length

Mean	1.7696
Standard Deviation	0.5106
Median	2
Minimum	1
Maximum	5

## Sample

1st row	7
2nd row	7
3rd row	4
4th row	4
5th row	8_

## Letter

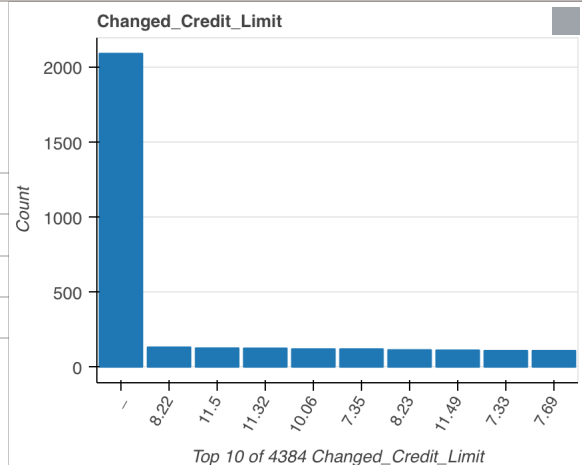
Count	0
Lowercase Letter	0
Space Separator	0
Uppercase Letter	0
Dash Punctuation	644
Decimal Number	161182

Changed\_Credit\_Li...

categorical

Hide Details

Approximate Distinct Count	4384
Approximate Unique (%)	4.4%
Missing	0
Missing (%)	0.0%
Memory Size	6.6 MB



Stats

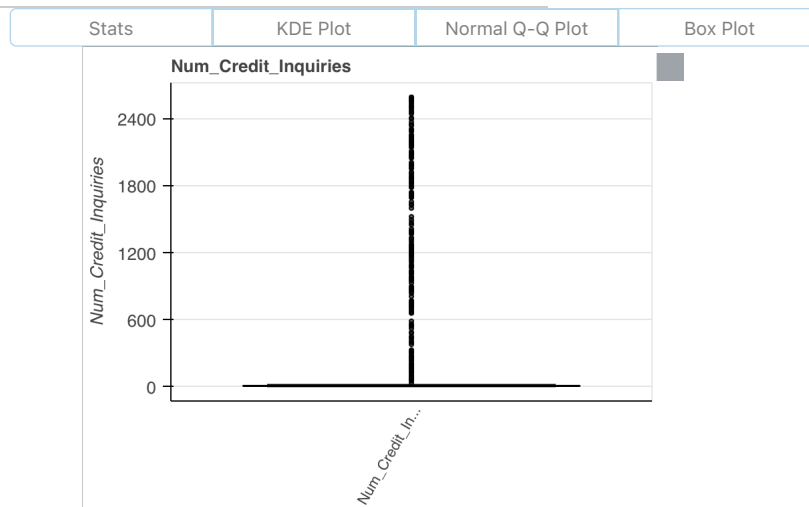
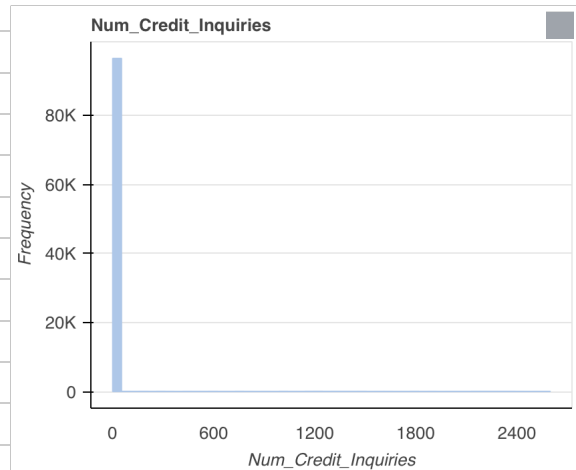
Pie Chart

Word Frequency

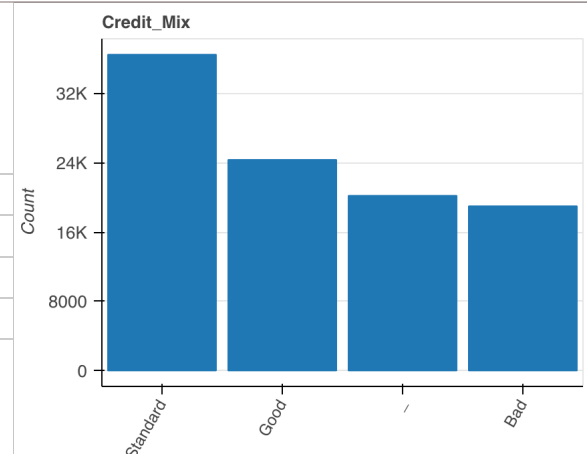
Word Length

Length		Sample		Letter	
Mean	4.7105	1st row	11.27	Count	0
Standard Deviation	2.4419	2nd row	11.27	Lowercase Letter	0
Median	4	3rd row	—	Space Separator	0
Minimum	1	4th row	6.27	Uppercase Letter	0
Maximum	21	5th row	11.27	Dash Punctuation	1586
				Decimal Number	369469

Num_Credit_Inquiri... numerical <a href="#">Hide Details</a>	Approximate Distinct Count	1223
	Approximate Unique (%)	1.2%
	Missing	1965
	Missing (%)	2.0%
	Infinite	0
	Infinite (%)	0.0%
	Memory Size	1.5 MB
	Mean	27.7543
	Minimum	0
	Maximum	2597
	Zeros	6972
	Zeros (%)	7.0%
	Negatives	0
	Negatives (%)	0.0%

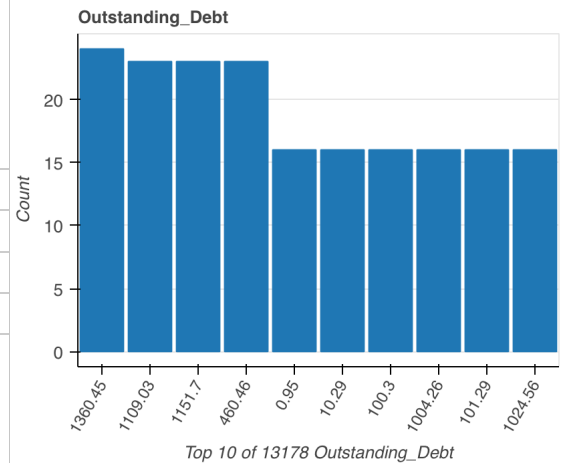


Credit_Mix categorical <a href="#">Hide Details</a>	Approximate Distinct Count	4
	Approximate Unique (%)	0.0%
	Missing	0
	Missing (%)	0.0%
	Memory Size	6.6 MB



Length		Sample		Letter	
Mean	4.6634	1st row	—	Count	446147
Standard Deviation	2.721	2nd row	Good	Lowercase Letter	366342
Median	4	3rd row	Good	Space Separator	0
Minimum	1	4th row	Good	Uppercase Letter	79805
Maximum	8	5th row	Good	Dash Punctuation	0
				Decimal Number	0

<b>Outstanding_Debt</b> categorical <a href="#">Hide Details</a>	Approximate Distinct Count	13178
	Approximate Unique (%)	13.2%
	Missing	0
	Missing (%)	0.0%
	Memory Size	6.8 MB



Stats

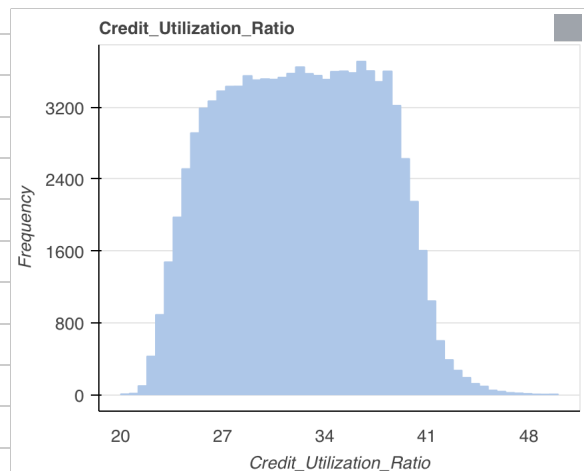
Pie Chart

Word Frequency

Word Length

Length		Sample		Letter	
Mean	6.4333	1st row	809.98	Count	0
Standard Deviation	0.6691	2nd row	809.98	Lowercase Letter	0
Median	7	3rd row	809.98	Space Separator	0
Minimum	3	4th row	809.98	Uppercase Letter	0
Maximum	8	5th row	809.98	Dash Punctuation	0
				Decimal Number	542320

<b>Credit_Utilization_...</b> numerical <a href="#">Hide Details</a>	Approximate Distinct Count	100000
	Approximate Unique (%)	100.0%
	Missing	0
	Missing (%)	0.0%
	Infinite	0
	Infinite (%)	0.0%
	Memory Size	1.5 MB
	Mean	32.2852
	Minimum	20
	Maximum	50
	Zeros	0
	Zeros (%)	0.0%
	Negatives	0
	Negatives (%)	0.0%

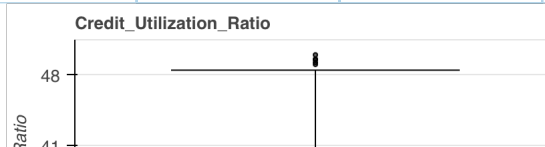


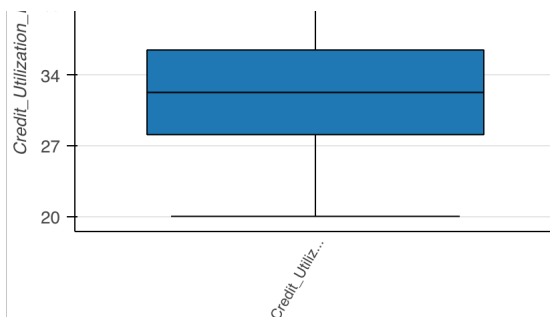
Stats

KDE Plot

Normal Q-Q Plot

Box Plot

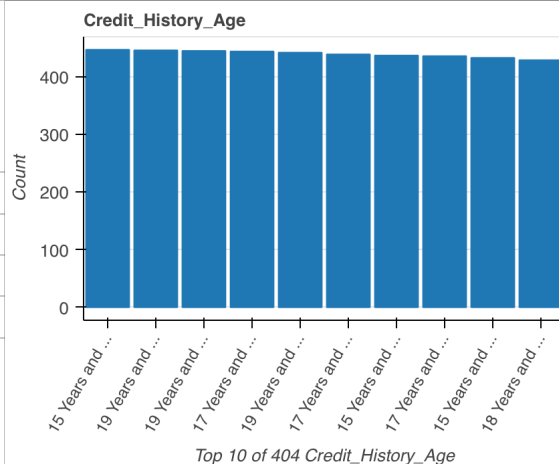


**Credit\_History\_Age**

categorical

[Hide Details](#)

Approximate Distinct Count	404
Approximate Unique (%)	0.4%
Missing	9030
Missing (%)	9.0%
Memory Size	7.5 MB



Stats

Pie Chart

Word Frequency

Word Length

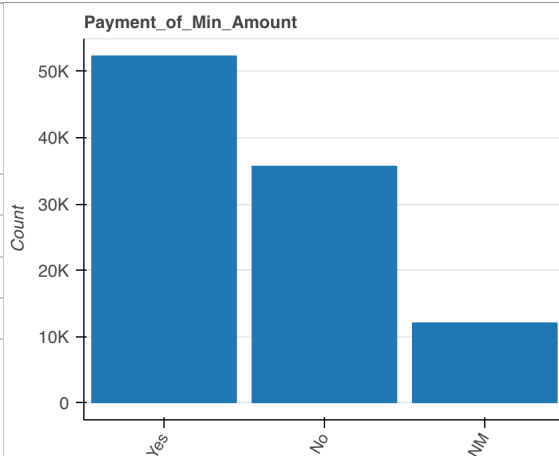
Length		Sample		Letter	
Mean	20.9867	1st row	22 Years and 1 Mon...	Count	1273580
Standard Deviation	0.5437	2nd row	22 Years and 3 Mon...	Lowercase Letter	1091640
Median	21	3rd row	22 Years and 4 Mon...	Space Separator	363880
Minimum	20	4th row	22 Years and 5 Mon...	Uppercase Letter	181940
Maximum	22	5th row	22 Years and 6 Mon...	Dash Punctuation	0
				Decimal Number	271703

**Payment\_of\_Min\_A...**

categorical

[Hide Details](#)

Approximate Distinct Count	3
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6.4 MB



Stats

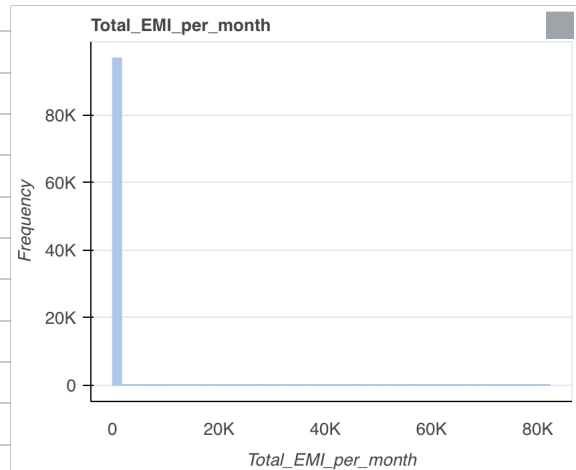
Pie Chart

Word Frequency

Word Length

Length		Sample		Letter	
Mean	2.5233	1st row	No	Count	252326
Standard Deviation	0.4995	2nd row	No	Lowercase Letter	140319
Median	3	3rd row	No	Space Separator	0
Minimum	2	4th row	No	Uppercase Letter	112007
Maximum	3	5th row	No	Dash Punctuation	0
				Decimal Number	0

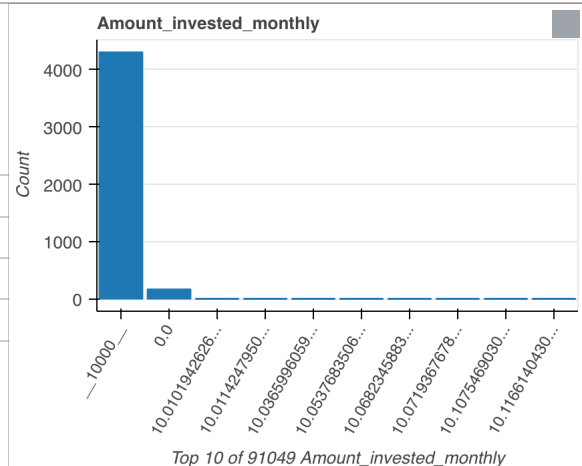
<b>Total_EMI_per_mo...</b> numerical <button>Hide Details</button>	Approximate Distinct Count	14950
	Approximate Unique (%)	14.9%
	Missing	0
	Missing (%)	0.0%
	Infinite	0
	Infinite (%)	0.0%
	Memory Size	1.5 MB
	Mean	1403.1182
	Minimum	0
	Maximum	82331
	Zeros	10613
	Zeros (%)	10.6%
	Negatives	0
	Negatives (%)	0.0%



[Stats](#) [KDE Plot](#) [Normal Q-Q Plot](#) [Box Plot](#)

Quantile Statistics		Descriptive Statistics	
Minimum	0	Mean	1403.1182
5-th Percentile	0	Standard Deviation	8306.0413
Q1	29.7564	Variance	6.899×10 <sup>07</sup>
Median	69.0433	Sum	1.4031×10 <sup>08</sup>
Q3	157.741	Skewness	7.1024
95-th Percentile	420.8887	Kurtosis	52.2176
Maximum	82331	Coefficient of Variation	5.9197
Range	82331		
IQR	127.9846		

<b>Amount_invested_...</b> categorical <button>Hide Details</button>	Approximate Distinct Count	91049
	Approximate Unique (%)	95.3%
	Missing	4479
	Missing (%)	4.5%
Memory Size		7.5 MB



[Stats](#) [Pie Chart](#) [Word Frequency](#) [Word Length](#)

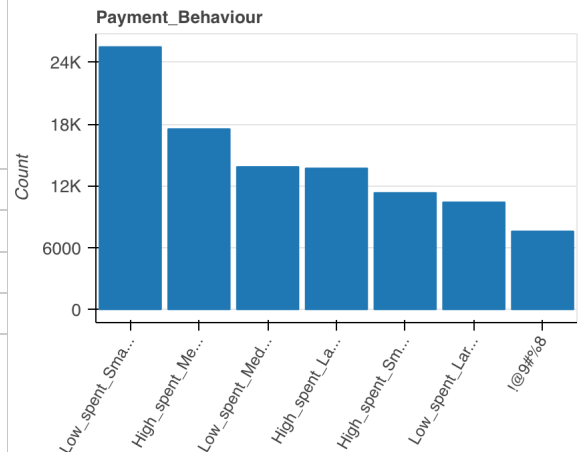
Length		Sample		Letter	
Mean	16.9655	1st row	80.41529543900253	Count	0
Standard Deviation	1.9337	2nd row	118.28022162236736	Lowercase Letter	0
Median	17	3rd row	81.699521264648	Space Separator	0
Minimum	3	4th row	199.4580743910713	Uppercase Letter	0
Maximum	18	5th row	41.420153086217326	Dash Punctuation	0
				Decimal Number	1512129

#### Payment\_Behaviour

categorical

[Hide Details](#)

Approximate Distinct Count	7
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	9.0 MB



Stats

Pie Chart

Word Frequency

Word Length

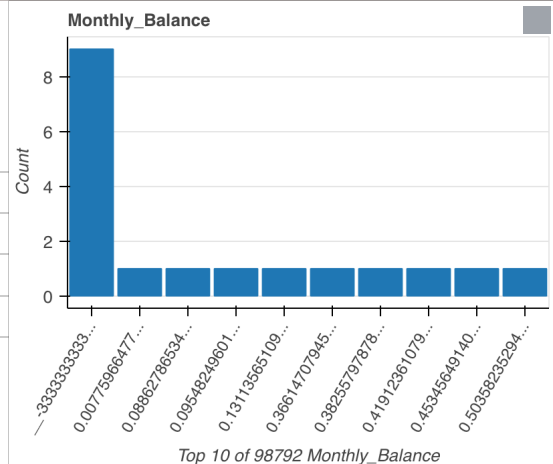
Length		Sample		Letter	
Mean	28.916	1st row	High_spent_Small_v...	Count	2476402
Standard Deviation	6.61	2nd row	Low_spent_Large_va...	Lowercase Letter	2291602
Median	31	3rd row	Low_spent_Medium_v...	Space Separator	0
Minimum	6	4th row	Low_spent_Small_va...	Uppercase Letter	184800
Maximum	32	5th row	High_spent_Medium_...	Dash Punctuation	0
				Decimal Number	15200

#### Monthly\_Balance

categorical

[Hide Details](#)

Approximate Distinct Count	98792
Approximate Unique (%)	100.0%
Missing	1200
Missing (%)	1.2%
Memory Size	7.8 MB



Stats

Pie Chart

Word Frequency

Word Length

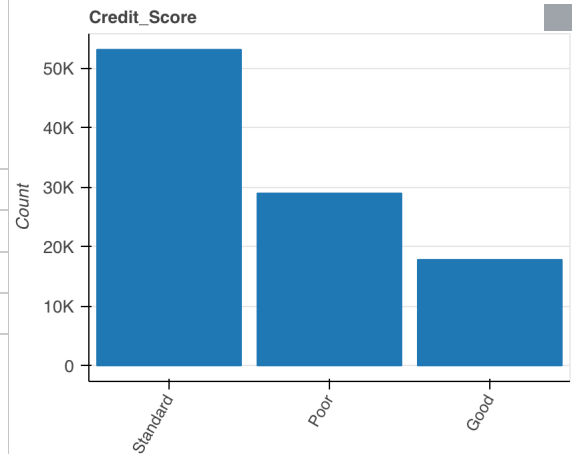
Length		Sample		Letter	
Mean	17.3407	1st row	312.49408867943663	Count	0
Standard Deviation	0.6273	2nd row	284.62916249607184	Lowercase Letter	0
Median	17	3rd row	331.2098628537912	Space Separator	0
Minimum	12	4th row	223.45130972736786	Uppercase Letter	0
Maximum	32	5th row	341.48923103222177	Dash Punctuation	9
				Decimal Number	1614427

### Credit\_Score

categorical

[Hide Details](#)

Approximate Distinct Count	3
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6.8 MB



Stats

Pie Chart

Word Frequency

Word Length

Length		Sample		Letter	
Mean	6.127	1st row	Good	Count	612696
Standard Deviation	1.996	2nd row	Good	Lowercase Letter	512696
Median	8	3rd row	Good	Space Separator	0
Minimum	4	4th row	Good	Uppercase Letter	100000
Maximum	8	5th row	Good	Dash Punctuation	0
				Decimal Number	0

## Interactions

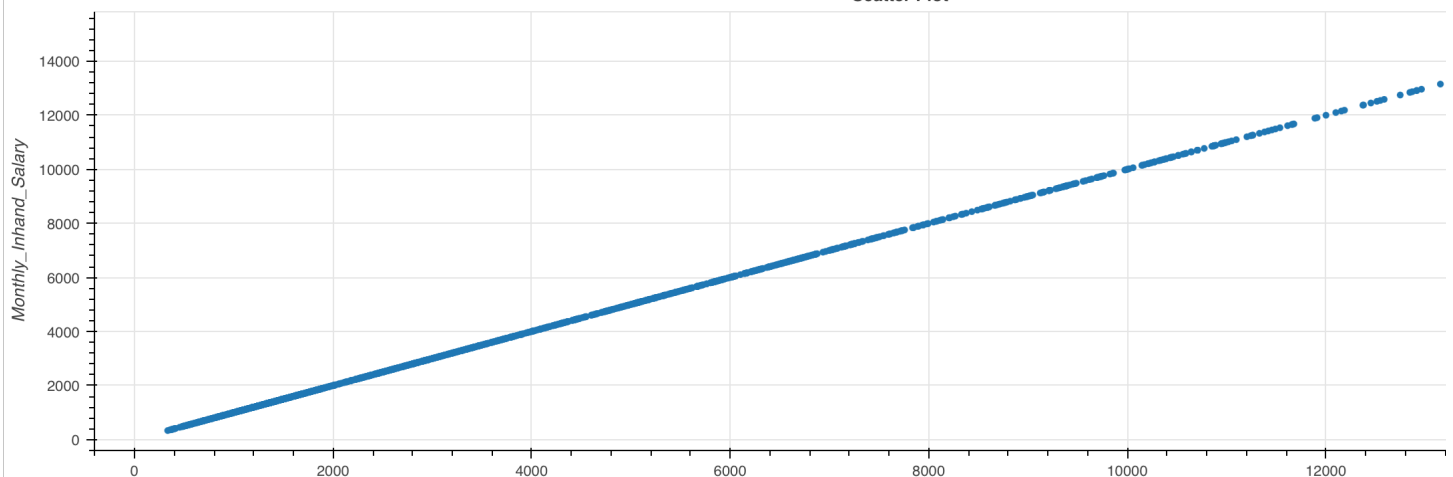
X-Axis

Monthly\_Inhand\_S

Y-Axis

Monthly\_Inhand\_S

Scatter Plot



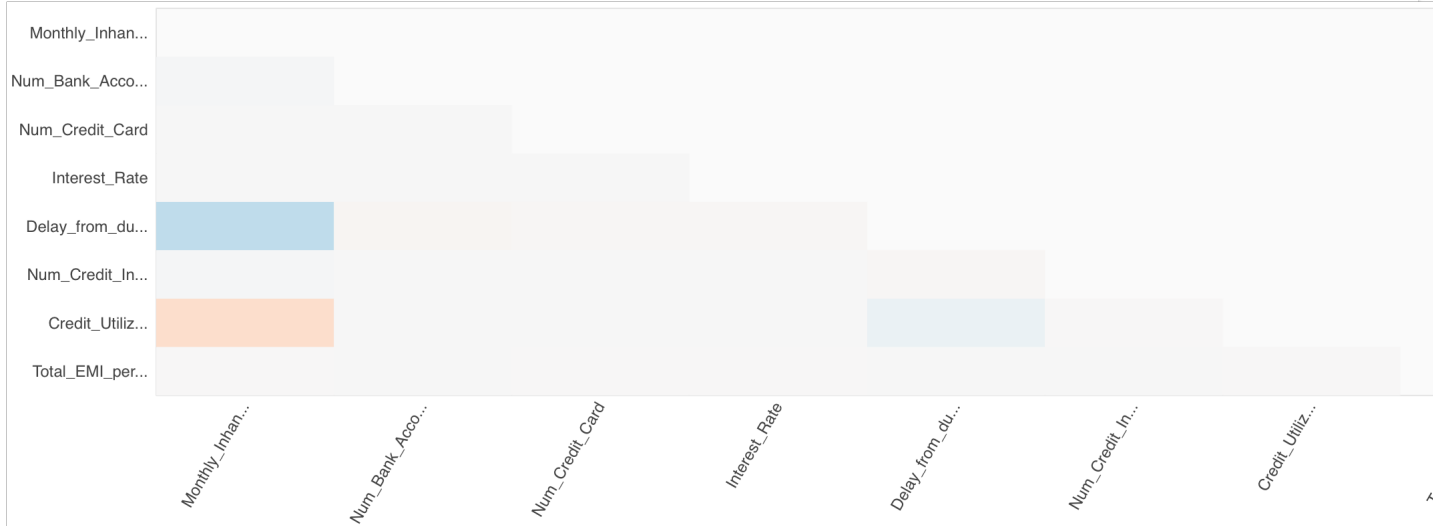
Monthly\_Inhand\_Salary

## Correlations

Pearson

Spearman

KendallTau



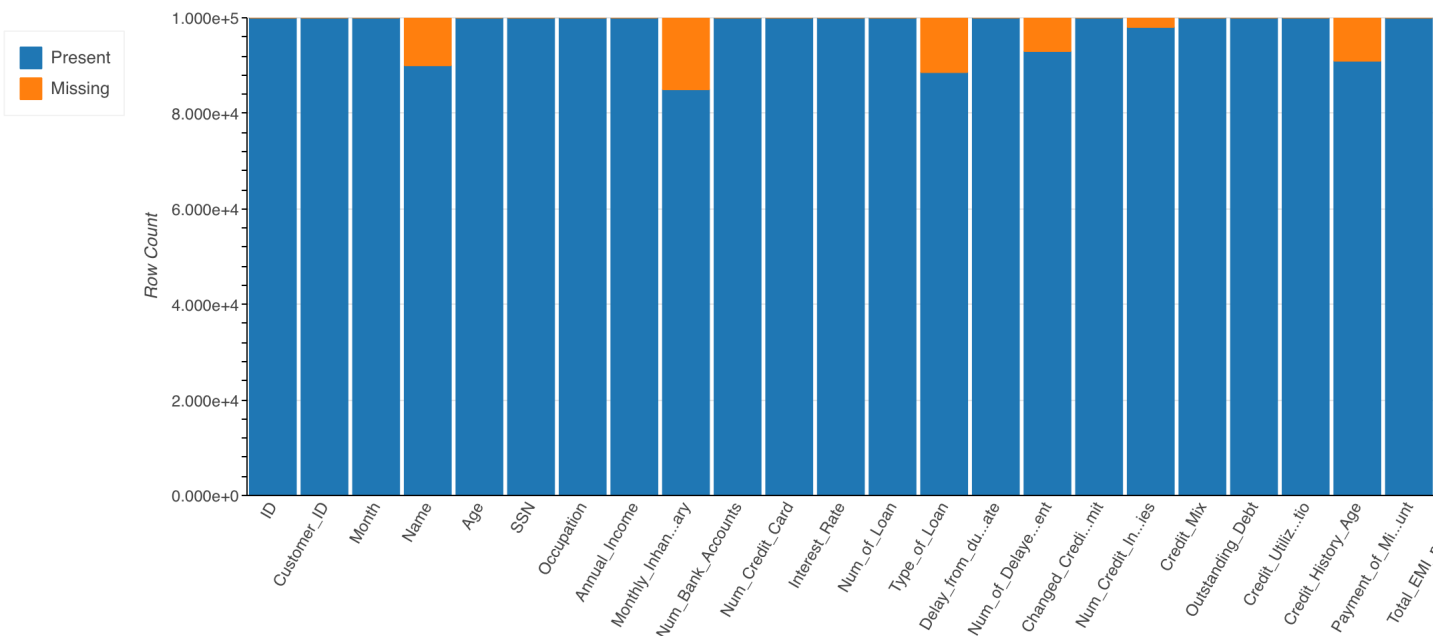
## Missing Values

Bar Chart

Spectrum

Heat Map

Dendrogram

Report generated with [DataPrep](#)