

Input Dataset

person_age	person_income	loan_intent	loan_amnt	loan_status	paid_past_due	person_home_ownership	person_emp_length	loan_grade	loan_int_rate	...
46	59000	VENTURE	35000	0	0	OWN	22	A	7.1	...
18	13000	EDUCATION	20000	1	110	OWN	13	A	8.3	...
20	24000	PERSONAL	30000	0	3	RENT	3	C	11.8	...
22	10000	EDUCATION	7900	1	93	RENT	5	D	12.9	...
35	36000	MEDICAL	50000	0	40	RENT	20	B	7.3	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	18	B	7.2	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	21	A	6.9	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	21	A	8.2	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	18	B	8.0	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	15	B	9.1	...

Input Dataset

Number of rows: 32581

Number of columns: 13

person_age	person_income	loan_intent	loan_amnt	loan_status	paid_past_due	person_home_ownership	person_emp_length	loan_grade	loan_int_rate	...
46	59000	VENTURE	35000	0	0	OWN	22	A	7.1	...
18	13000	EDUCATION	20000	1	110	OWN	13	A	8.3	...
20	24000	PERSONAL	30000	0	3	RENT	3	C	11.8	...
22	10000	EDUCATION	7900	1	93	RENT	5	D	12.9	...
35	36000	MEDICAL	50000	0	40	RENT	20	B	7.3	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	18	B	7.2	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	21	A	6.9	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	21	A	8.2	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	18	B	8.0	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	15	B	9.1	...

Define Good & Bad Definitions

paid_past_due ✓

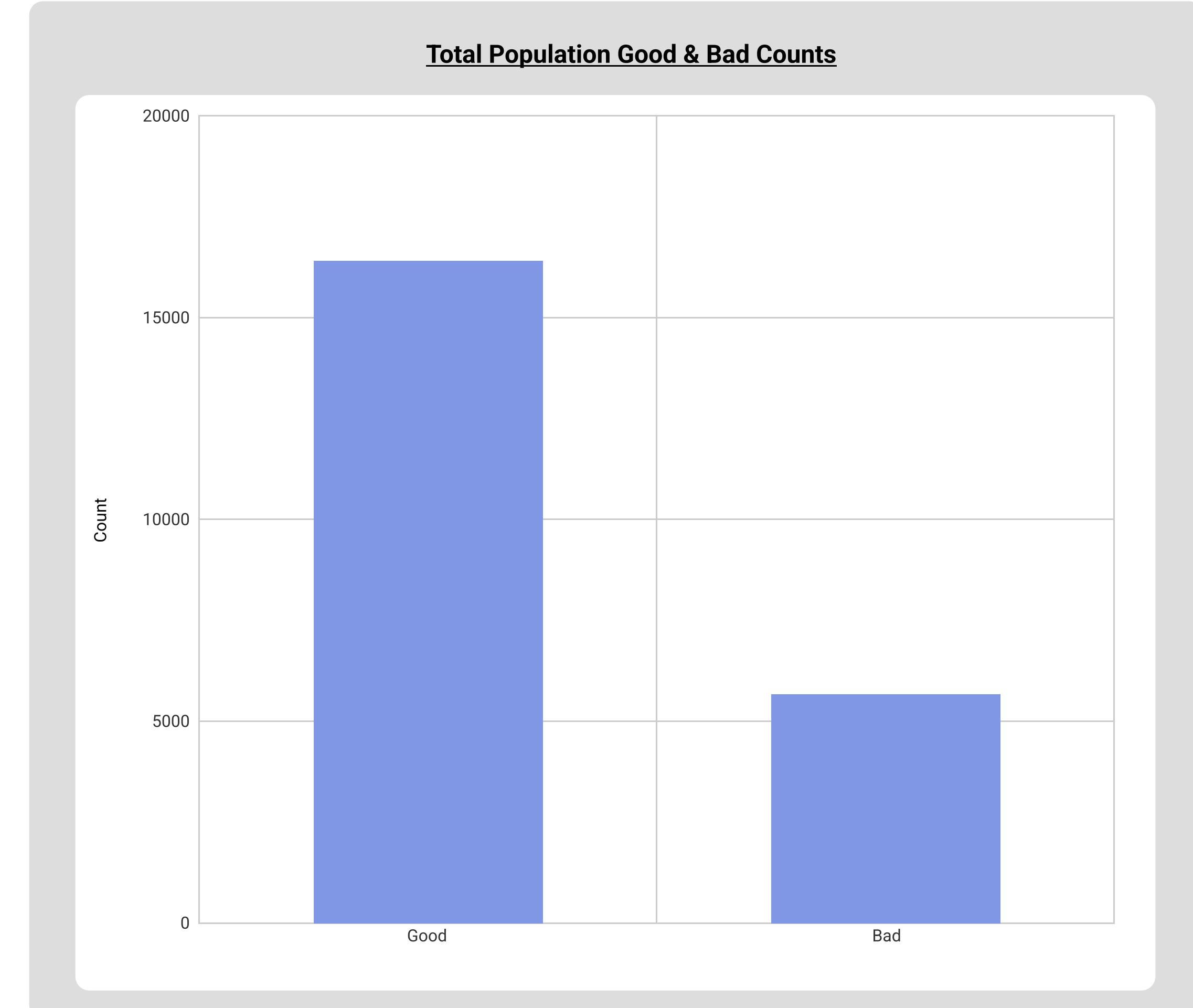
Weight of good: 1.49

Weight of bad: 1

0 120

*Select the range of indeterminate

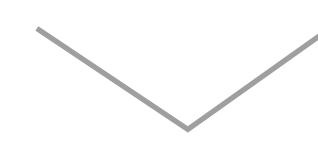
Refresh



Confirm Input Dataset

I. Select the columns you would like to bin:

x person_age	x person_income	x person_home_ownership	x person_emp_length	
x loan_status	x paid_past_due	x loan_grade	x loan_amnt	x loan_int_rate
x cb_person_default_on_file				



II. Preview your dataset here:

person_age	person_income	loan_intent	loan_amnt	loan_status	paid_past_due	person_home_ownership	person_emp_length	loan_grade	loan_int_rate	...
46	59000	VENTURE	35000	0	0	OWN	22	A	7.1	...
18	13000	EDUCATION	20000	1	110	OWN	13	A	8.3	...
20	24000	PERSONAL	30000	0	3	RENT	3	C	11.8	...
22	10000	EDUCATION	7900	1	93	RENT	5	D	12.9	...
35	36000	MEDICAL	50000	0	40	RENT	20	B	7.3	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	18	B	7.2	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	21	A	6.9	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	21	A	8.2	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	18	B	8.0	...
35	36000	MEDICAL	50000	0	40	MORTGAGE	15	B	9.1	...

« ‹ 1 / 3259 › »

III. Confirm the predictor type of the columns:

person_age	numerical
person_income	numerical
person_home_ownership	categorical
person_emp_length	numerical
loan_status	categorical
paid_past_due	numerical
loan_grade	categorical
loan_amnt	numerical
loan_int_rate	numerical
cb_person_default_on_file	categorical

IV. Save & Confirm your settings:

Save & Confirm Settings

Define Good & Bad Definitions

I. Define Bad Definition

- Numerical Variable

1. person_age : 0 (inclusive) - 20 (exclusive)

2. paid_past_due : 90 (inclusive) - 120 (exclusive)

AddRemove

- Categorical Variable

1. person_home_ow... : **x MORTGAGE** **x RENT**

2. loan_status : **x 1**

AddRemove

Weight of bad:

1

II. Define Indeterminate Definition

- Numerical Variable

1. person_age : 20 (inclusive) - 30 (exclusive)

2. paid_past_due : 60 (inclusive) - 90 (exclusive)

AddRemove

- Categorical Variable

1. person_home_ow... : **x OTHER**

AddRemove

III. Define Good Definition

Weight of good:

1

IV. Confirm the Definitions:

Confirm

Interactive Binning

I. Select a predictor variable to bin

person_age



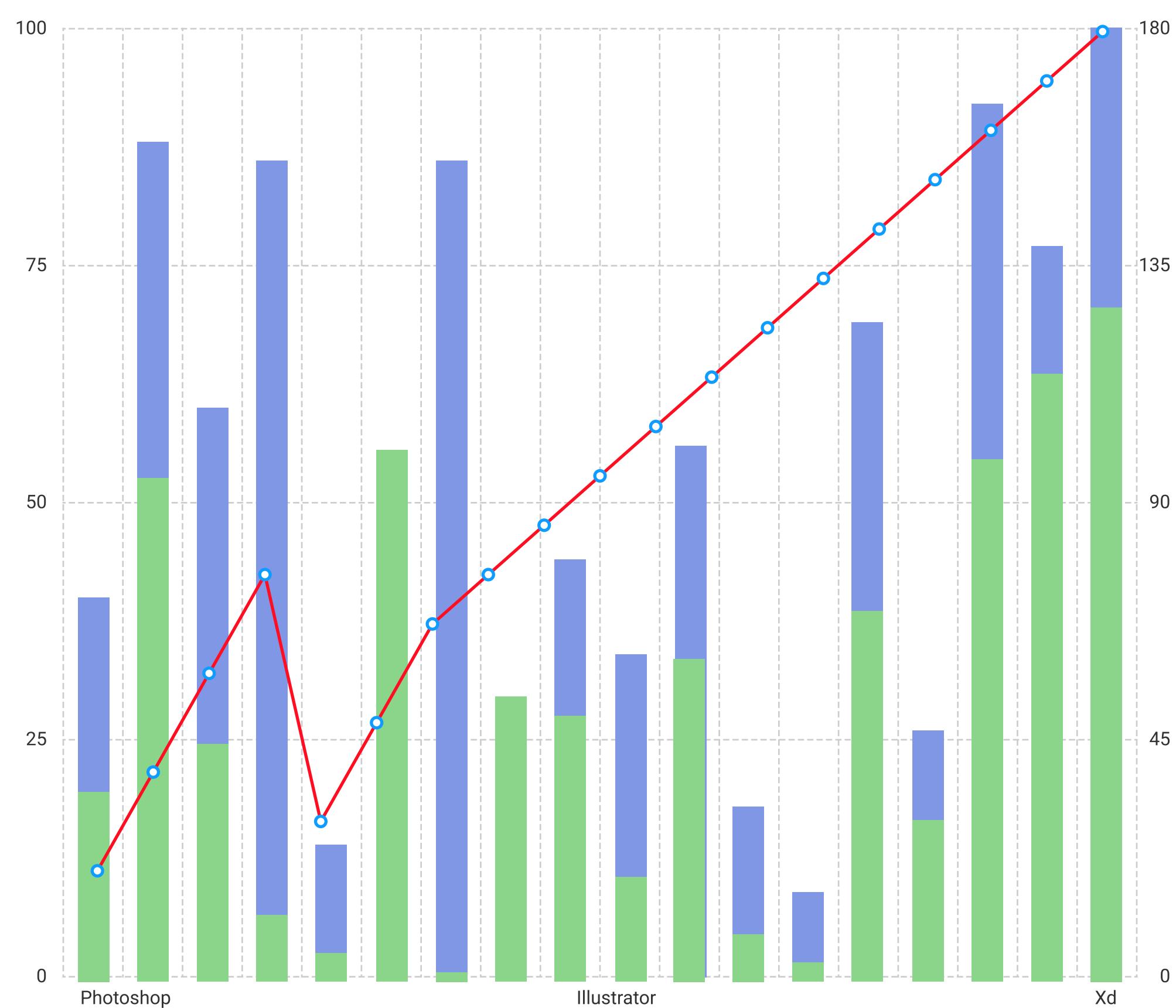
II. Select the automated binning algorithm for initial binning

No Binnings

**Refresh**

*Regards each unique value in the dataset as a bin

III. Perform Interactive Binning



Selected Bin Name: [[0, 20), [80, 90]]

Population Good Count: 2391

Population Bad Count: 193

Show Details

IV. Monitor Bins Performance (Before)

Bin	Good	Bad	Odds	Total	Good%	Bad%	Total%	Info_Odds	WOE	MC
RENT	11254	5192	2.1676	16446	44.18	73.04	50.48	0.6048	-0.5028	0.1451
OWN	2391	193	12.3886	2584	9.39	2.72	7.93	3.4569	1.2404	0.0827
MORTGAGE	11754	1690	6.9550	13444	46.14	23.78	41.26	1.9407	0.6631	0.1483
OTHER	74	33	2.2424	107	0.29	0.46	0.33	0.6257	-0.4688	0.0000
Total	25473	7108	3.5837	32581	100.00	100.00	100.00			0.3770

V. Monitor Bins Performance (After)

Bin	Good	Bad	Odds	Total	Good%	Bad%	Total%	Info_Odds	WOE	MC
RENT	11254	5192	2.1676	16446	44.18	73.04	50.48	0.6048	-0.5028	0.1451
['OWN', 'OTHERS']	2391	193	12.3886	2584	9.39	2.72	7.93	3.4569	1.2404	0.0827
MORTGAGE	11754	1690	6.9550	13444	46.14	23.78	41.26	1.9407	0.6631	0.1483
Total	25473	7108	3.5837	32581	100.00	100.00	100.00			0.3770

VI. Save & Confirm Your Bins Settings for the Chosen Predictor Variable:

Confirm Binning

II. Select the automated binning algorithm for initial binning

Equal Width 

Width Number of Bins

Width: 1

Refresh

*Divides the range of value with predetermined width OR into predetermined number of equal width bins

Selected Bin Name: ['RENT', 'MORTGAGE']

Population Good Count: 11754

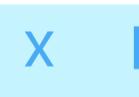
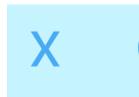
Population Bad Count: 1690

Rename Bin To: Risky **Submit**

Usage 1: Split the bin into two by indicating the element(s) to be included in one of the bin

Usage 2: Add elements from other bin(s) to the selected bin, the added element(s) will be automatically removed from the other bin(s)

Bin Element(s):

 RENT  OWN 

Add

Preview Changes

Old Bin(s):

(1) Old Bin Name: ['RENT', 'MORTGAGE']

Old Bin Elements: ['RENT', 'MORTGAGE']

(2) Old Bin Name: Others

Old Bin Element(s): ['OWN', 'OTHER']

Will be changed to:

(1) New Bin Name: ['RENT', 'OWN']

New Bin Element(s): ['RENT', 'OWN']

(2) New Bin Name: ['MORTGAGE']

New Bin Element(s): ['MORTGAGE']

(3) New Bin Name: Others

New Bin Element(s): ['OTHER']

Submit

Hide Details

*Note: Submitting the changes only updates the mixed chart & the statistical tables, it DOES NOT save the bins settings until you click the 'Confirm Binning' button in Section V.

II. Select the automated binning algorithm for initial binning

Equal Frequency 

Frequency Number of Bins

Frequency: 1000

Refresh

*Divides the data into a predetermined number of bins containing approximately the same number of observations

Selected Bin Name: [[0, 20), [80, 90]]

Population Good Count: 2391

Population Bad Count: 193

Rename Bin To: Risky **Submit**

Usage 1: Narrow down ranges to split bin into two

Usage 2: Amend bin boundaries (conflicting boundaries of another bin(s) will be automatically narrowed down

Bin Range(s):

1. 0 - 10
2. 80 - 110

Add

Preview Changes

Old Bin(s):

(1) Old Bin Name: [[0, 20), [80, 90]]

Old Bin Ranges: [[0, 20), [80, 90]]

(2) Old Bin Name: Elderly

Old Bin Ranges: [[90, 120]]

Will be changed to:

(1) New Bin Name: [[0, 10), [80, 110]]

New Bin Ranges: [[0, 10), [80, 110]]

(2) New Bin Name: [[10, 20]]

New Bin Ranges: [[10, 20]]

(3) New Bin Name: Elderly

New Bin Ranges: [[110, 120]]

Submit

Hide Details

*Note: Submitting the changes only updates the mixed chart & the statistical tables, it DOES NOT save the bins settings until you click the 'Confirm Binning' button in Section V.

Selected Bin Name: [[0, 20), [80, 90]]

Population Good Count: 11754

Population Bad Count: 1690

Bin Range(s): [[0, 20), [80, 90]]

Selected Bin Name: Elderly

Population Good Count: 74

Population Bad Count: 33

Bin Range(s): [[90, 120]]

Selected Bin Name: [[30, 35), [40, 45]]

Population Good Count: 2391

Population Bad Count: 193

Bin Range(s): [[30, 35), [40, 45]]

Preview Changes

New Bin:

Bin Name: [[0, 20), [30, 35), [40, 45), [80, 120]]

Bin Elements: [[0, 20), [30, 35), [40, 45), [80, 120]]

Submit

Hide Details

*Note: Submitting the changes only updates the mixed chart & the statistical tables, it DOES NOT save the bins settings until you click the 'Confirm Binning' button in Section V.

Show Details

Selected Bin Name: ['RENT', 'MORTGAGE']

Population Good Count: 11754

Population Bad Count: 1690

Bin Element(s): ['RENT', 'MORTGAGE']

Selected Bin Name: Other

Population Good Count: 74

Population Bad Count: 33

Bin Element(s): ['OTHER']

Selected Bin Name: ['OWN']

Population Good Count: 2391

Population Bad Count: 193

Bin Element(s): ['OWN']

Preview Changes

New Bin:

Bin Name: ['RENT', 'MORTGAGE', 'OTHER', 'OWN']

Bin Element(s): ['RENT', 'MORTGAGE', 'OTHER', 'OWN']

Submit

Hide Details

*Note: Submitting the changes only updates the mixed chart & the statistical tables, it DOES NOT save the bins settings until you click the 'Confirm Binning' button in Section V.

Preview & Download Settings

I. Preview Output Dataset

person_age	person_income	loan_intent	loan_amnt	loan_status	paid_past_due	person_age_bin_ned	person_income_binned	loan_intent_bin_ned	loan_amnt_bin_ned	...
46	59000	VENTURE	35000	0	0	Mid-Age	High	VENTURE	Mid	...
18	13000	EDUCATION	20000	1	110	Young	Low	EDUCATION	Mid	...
20	24000	PERSONAL	30000	0	3	Young	Mid	PERSONAL	Mid	...
22	10000	EDUCATION	7900	1	93	Young	Low	EDUCATION	Low	...
35	36000	MEDICAL	50000	0	40	Mid-Age	Mid	MEDICAL	High	...
35	36000	MEDICAL	50000	0	40	Mid-Age	Mid	MEDICAL	High	...
35	36000	MEDICAL	50000	0	40	Mid-Age	Mid	MEDICAL	High	...
35	36000	MEDICAL	50000	0	40	Mid-Age	Mid	MEDICAL	High	...
35	36000	MEDICAL	50000	0	40	Mid-Age	Mid	MEDICAL	High	...
35	36000	MEDICAL	50000	0	40	Mid-Age	Mid	MEDICAL	High	...

<< < 1 / 3259 > >>

II. Preview Bins' Performance

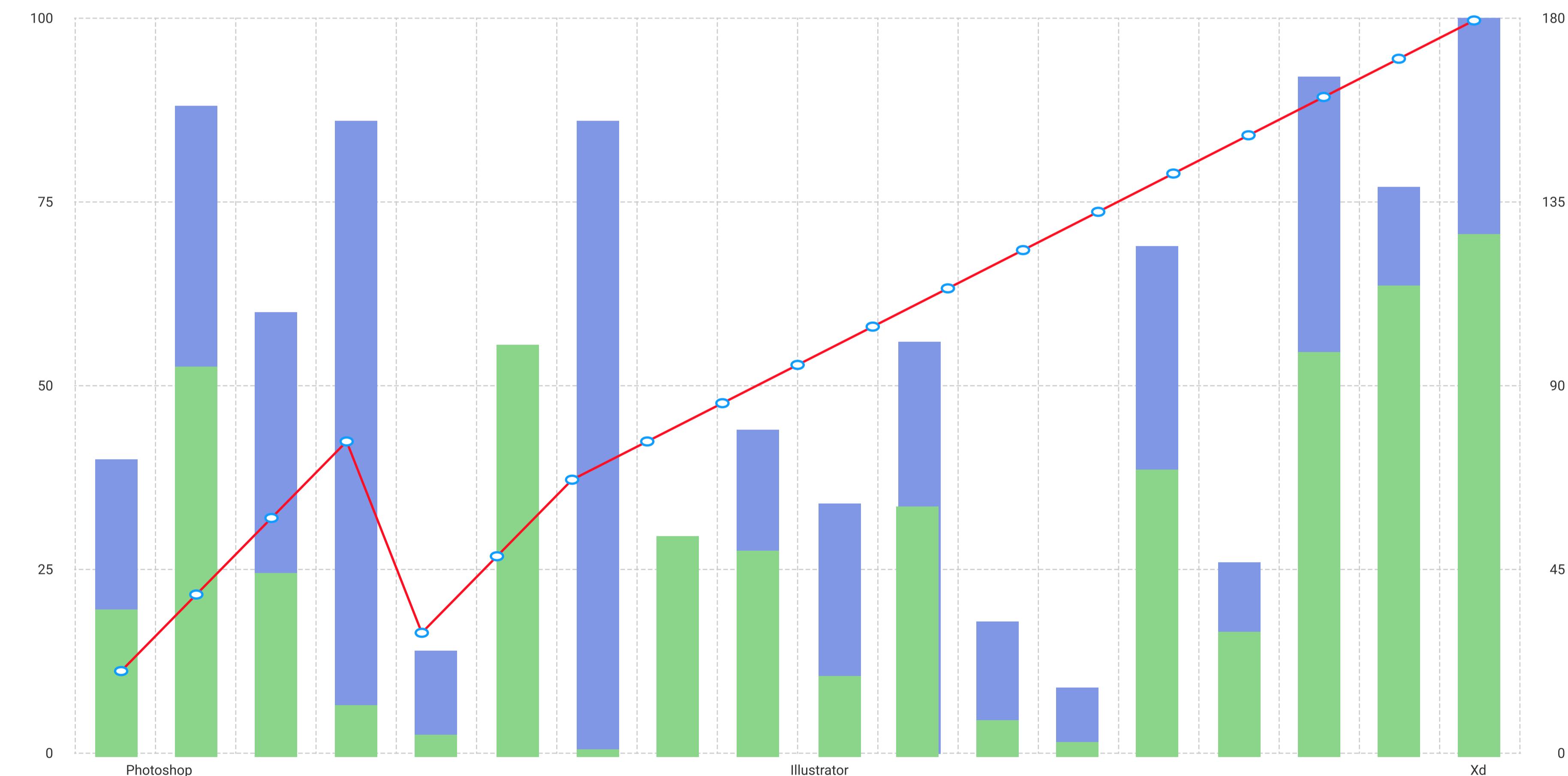
Select variable to preview:

person_home_ownership

Summary Statistical Table

Bin	Good	Bad	Odds	Total	Good%	Bad%	Total%	Info_Odds	WOE	MC
RENT	11254	5192	2.1676	16446	44.18	73.04	50.48	0.6048	-0.5028	0.1451
['OWN', 'OTHERS']	2391	193	12.3886	2584	9.39	2.72	7.93	3.4569	1.2404	0.0827
MORTGAGE	11754	1690	6.9550	13444	46.14	23.78	41.26	1.9407	0.6631	0.1483
Total	25473	7108	3.5837	32581	100.00	100.00	100.00			0.3770

Mixed Chart



III. Download Bins Settings

[Download Bins Settings](#)

After download the bins_settings.json file, you could replace it with the one in the Dataiku flow, re-run the flow, and bin the dataset there. You could see the same statistical table and mixed chart in the dashboard too.