

Overview

Dataset Statistics

Number of Variables	50
Number of Rows	101766
Missing Cells	192849
Missing Cells (%)	3.8%
Duplicate Rows	0
Duplicate Rows (%)	0.0%
Total Size in Memory	221.9 MB
Average Row Size in Memory	2.2 KB
Variable Types	Numerical: 11 Categorical: 39

Dataset Insights

[race]	has 2273 (2.23%) missing values	Missing
[weight]	has 98569 (96.86%) missing values	Missing
[payer_code]	has 40256 (39.56%) missing values	Missing
[medical_specialty]	has 49949 (49.08%) missing values	Missing
[diag_3]	has 1423 (1.4%) missing values	Missing
[patient_nbr]	is skewed	Skewed
[discharge_disposition_id]	is skewed	Skewed
[admission_source_id]	is skewed	Skewed
[time_in_hospital]	is skewed	Skewed
[number_outpatient]	is skewed	Skewed

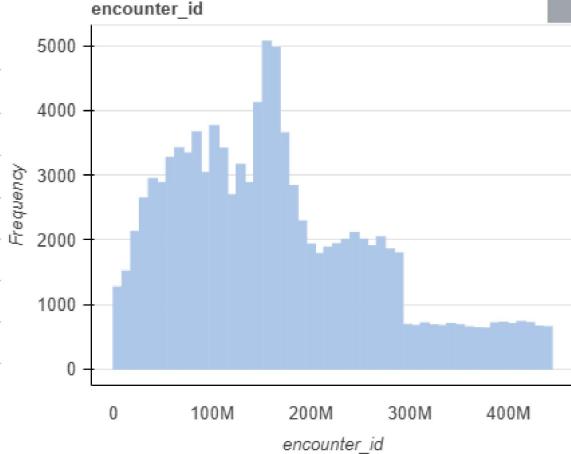
1 2 3

Variables

Sort by Feature order Reverse order

encounter_id numerical	Approximate Distinct Count	101766
	Approximate Unique (%)	100.0%
	Missing	0
	Missing (%)	0.0%
	Infinite	0
	Infinite (%)	0.0%
	Memory Size	1628256

Show Details



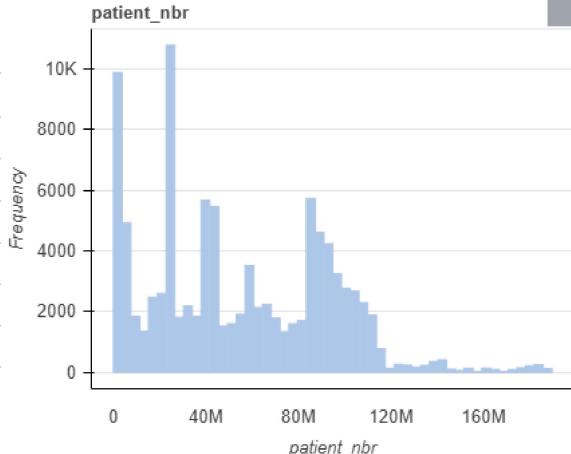
Frequency

encounter_id

0 100M 200M 300M 400M

patient_nbr numerical	Approximate Distinct Count	71518
	Approximate Unique (%)	70.3%
	Missing	0
	Missing (%)	0.0%
	Infinite	0
	Infinite (%)	0.0%
	Memory Size	1628256

Show Details



Frequency

patient_nbr

0 40M 80M 120M 160M

Approximate Distinct Count

5

Approximate Unique (%)

0.0%

Missing

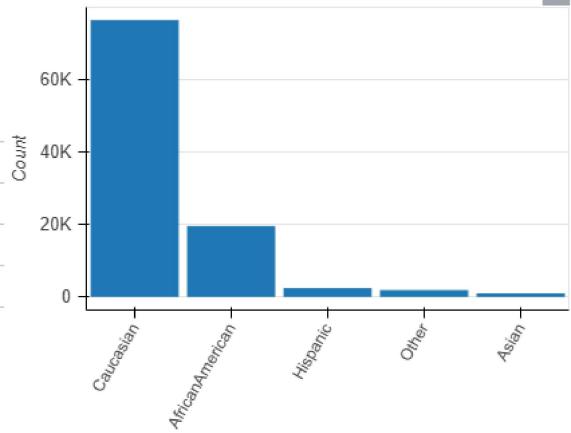
2273

Missing (%)

2.2%

Memory Size

7467117

race
 categorical
[Show Details](#)

Approximate Distinct Count

3

Approximate Unique (%)

0.0%

Missing

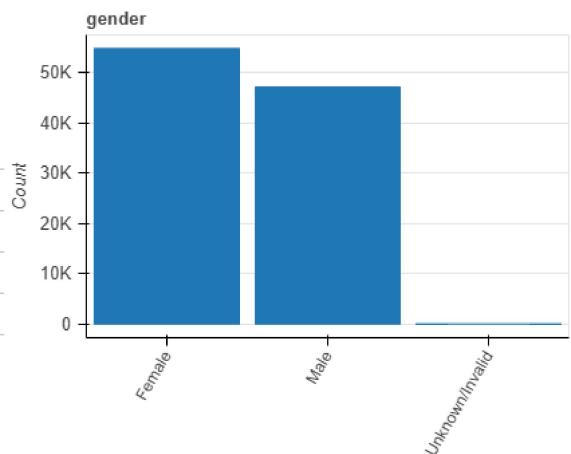
0

Missing (%)

0.0%

Memory Size

7131303

gender
 categorical
[Show Details](#)

Approximate Distinct Count

10

Approximate Unique (%)

0.0%

Missing

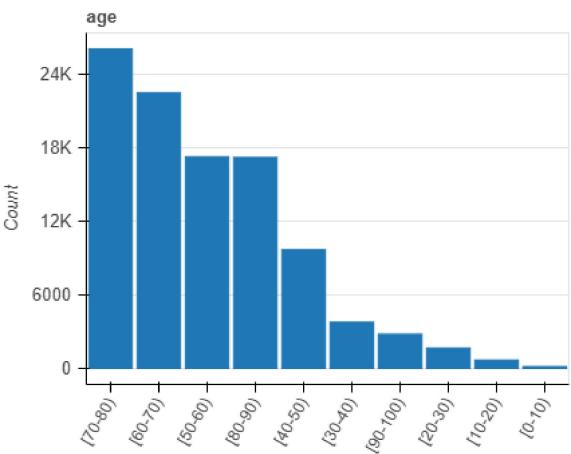
0

Missing (%)

0.0%

Memory Size

7329784

age
 categorical
[Show Details](#)

Approximate Distinct Count

9

Approximate Unique (%)

0.3%

Missing

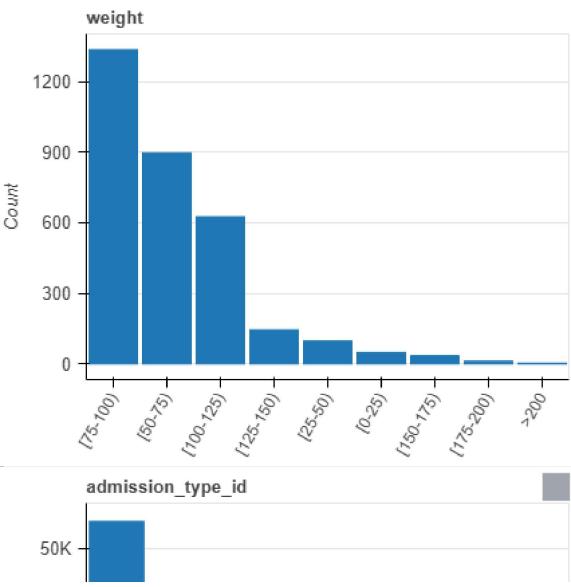
98569

Missing (%)

96.9%

Memory Size

233095

weight
 categorical
[Show Details](#)

DataPrep Report

Approximate Distinct Count

8

admission_type_id

categorical

[Show Details](#)

Approximate Unique (%)

0.0%

Missing

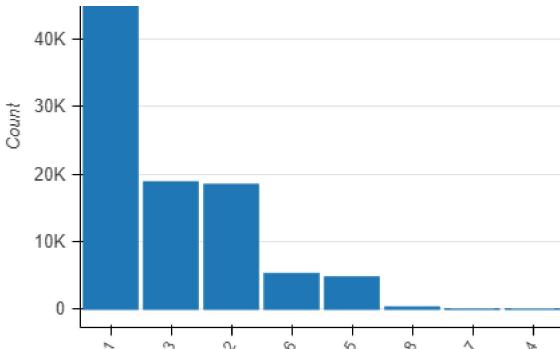
0

Missing (%)

0.0%

Memory Size

6716556

**discharge_dispo...**

numerical

[Show Details](#)

Approximate Distinct Count

26

Mean

3.7156

Approximate Unique (%)

0.0%

Minimum

1

Maximum

28

Zeros

0

Zeros (%)

0.0%

Negatives

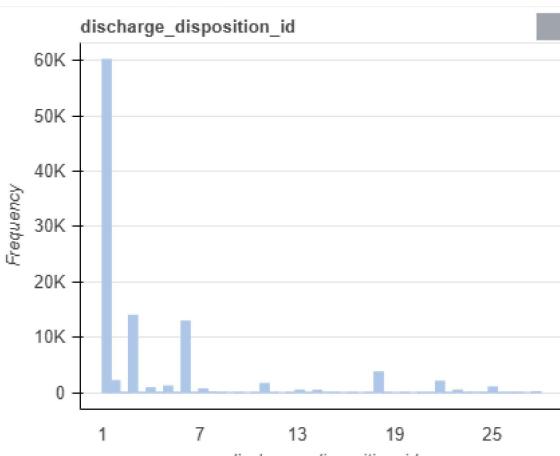
0

Negatives (%)

0.0%

Memory Size

1628256

**admission_source...**

numerical

[Show Details](#)

Approximate Distinct Count

17

Mean

5.7544

Approximate Unique (%)

0.0%

Minimum

1

Maximum

25

Zeros

0

Zeros (%)

0.0%

Negatives

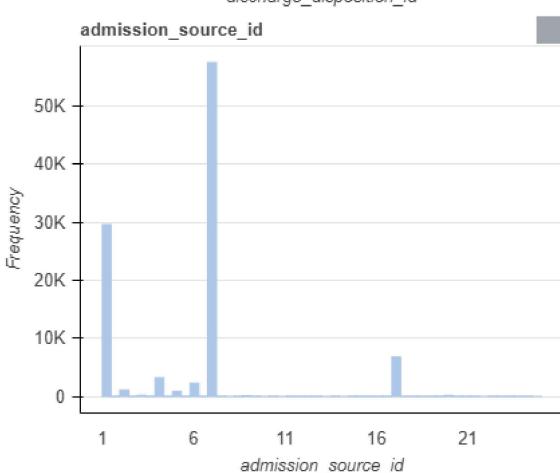
0

Negatives (%)

0.0%

Memory Size

1628256

**time_in_hospital**

numerical

[Show Details](#)

Approximate Distinct Count

14

Mean

4.396

Approximate Unique (%)

0.0%

Missing

0

Missing (%)

0.0%

Infinite

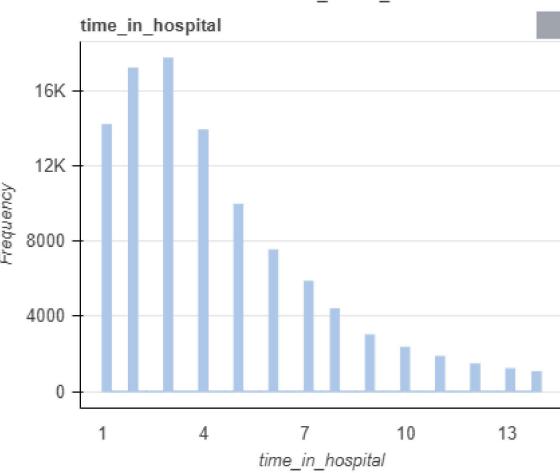
0

Infinite (%)

0.0%

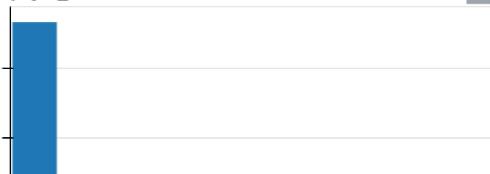
Memory Size

1628256



Approximate Distinct Count

17

payer_code

payer_code
categorical

Approximate Unique (%)	0.0%
Missing	40256
Missing (%)	39.6%
Memory Size	4121170

Show Details

Top 10 of 17 payer_code

medical_specialty
categorical

Approximate Distinct Count	72
Approximate Unique (%)	0.1%
Missing	49949
Missing (%)	49.1%
Memory Size	4194633

Show Details

Top 10 of 72 medical_specialty

num_lab_procedures
numerical

Approximate Distinct Count	118
Approximate Unique (%)	0.1%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Memory Size	1628256

Show Details

Frequency

num_lab_procedures

num_procedures
categorical

Approximate Distinct Count	7
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6716556

Show Details

Count

num_procedures

num_medications
numerical

Approximate Distinct Count	75
Approximate Unique (%)	0.1%
Missing	0
Missing (%)	0.0%

Show Details

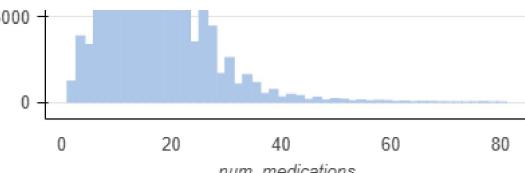
Frequency

num_medications

DataPrep Report

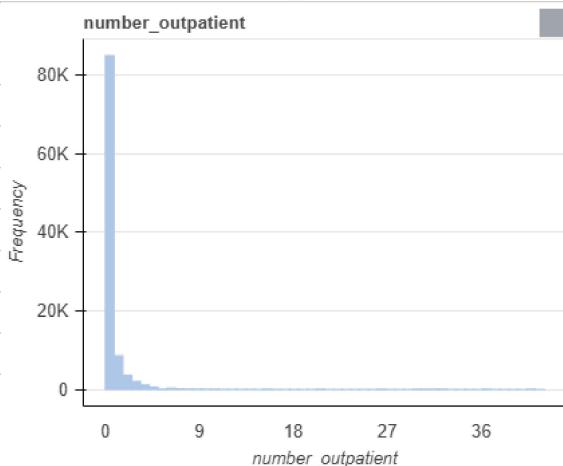
Maximum	81
Zeros	0
Zeros (%)	0.0%

Negatives	0
Negatives (%)	0.0%



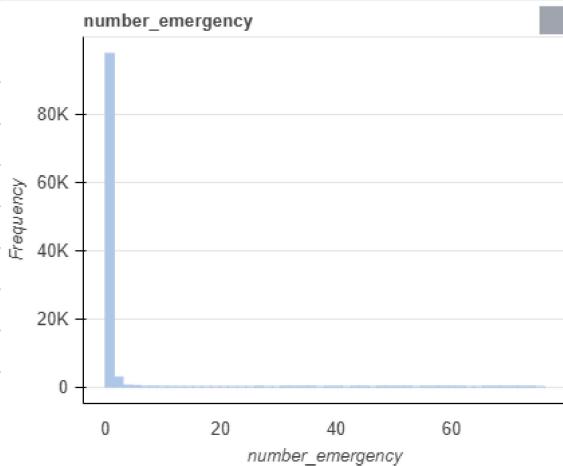
Approximate Distinct Count	39
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Memory Size	1628256

Mean	0.3694
Minimum	0
Maximum	42
Zeros	85027
Zeros (%)	83.5%
Negatives	0
Negatives (%)	0.0%



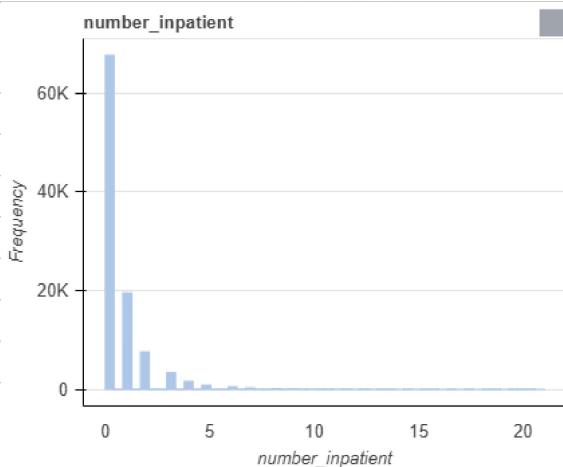
Approximate Distinct Count	33
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Memory Size	1628256

Mean	0.1978
Minimum	0
Maximum	76
Zeros	90383
Zeros (%)	88.8%
Negatives	0
Negatives (%)	0.0%

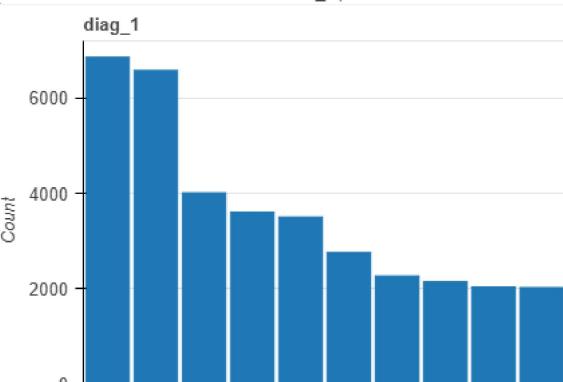


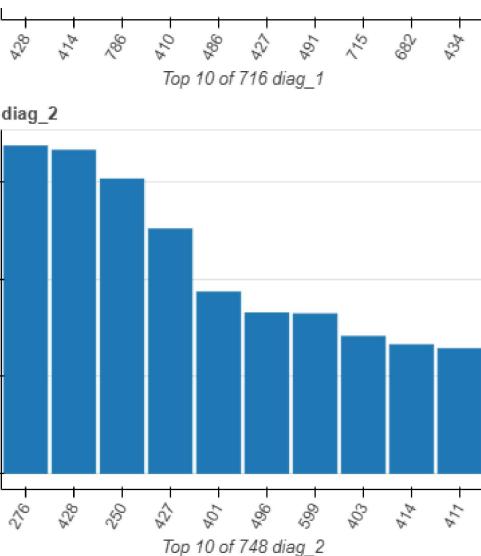
Approximate Distinct Count	21
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Memory Size	1628256

Mean	0.6356
Minimum	0
Maximum	21
Zeros	67630
Zeros (%)	66.5%
Negatives	0
Negatives (%)	0.0%



Approximate Distinct Count	716
Approximate Unique (%)	0.7%
Missing	21
Missing (%)	0.0%
Memory Size	6936533





diag_2
categorical

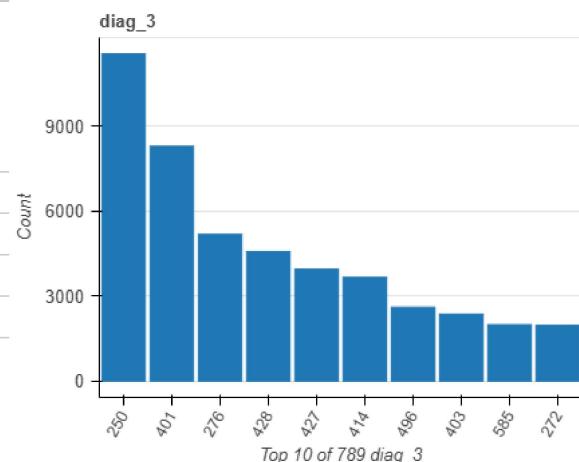
Show Details

Approximate Distinct Count	748
Approximate Unique (%)	0.7%
Missing	358
Missing (%)	0.4%
Memory Size	6913373

diag_3
categorical

Show Details

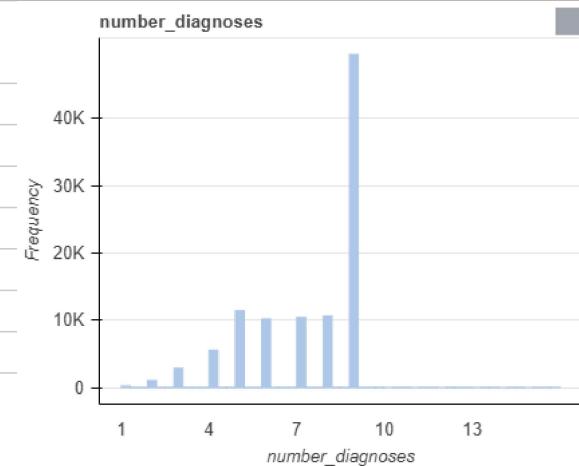
Approximate Distinct Count	789
Approximate Unique (%)	0.8%
Missing	1423
Missing (%)	1.4%
Memory Size	6837533



number_diagnoses
numerical

Show Details

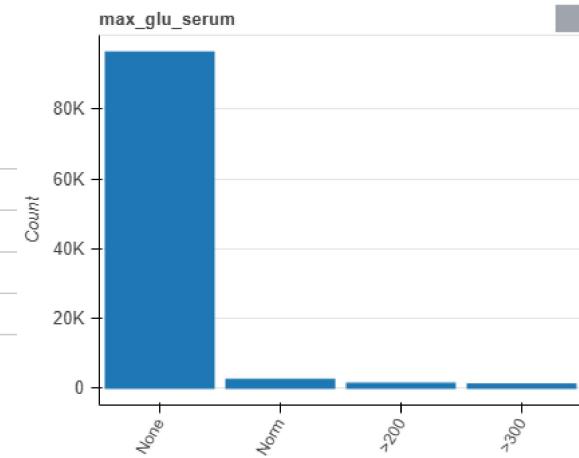
Approximate Distinct Count	16
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Infinite	0
Infinite (%)	0.0%
Memory Size	1628256

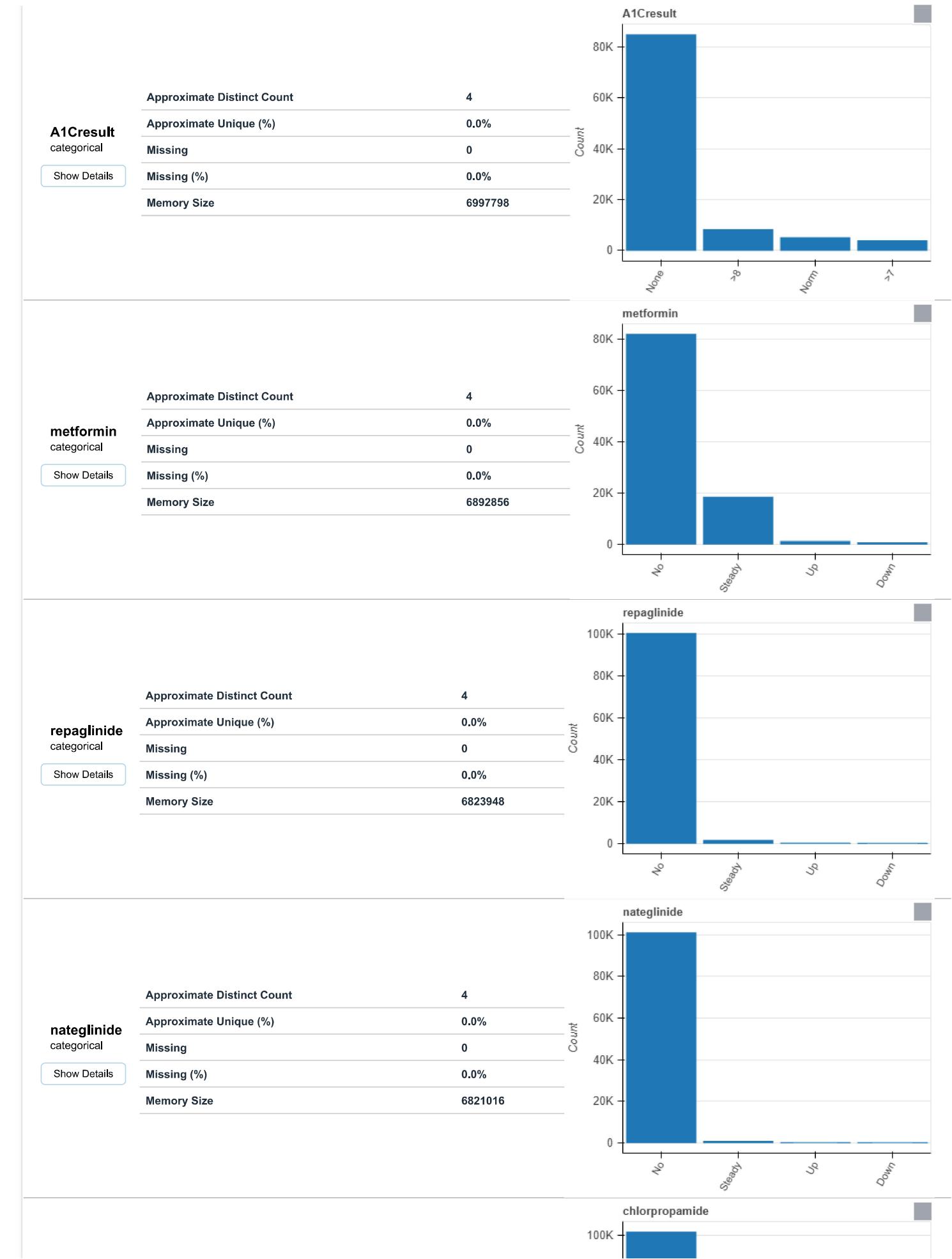


max_glu_serum
categorical

Show Details

Approximate Distinct Count	4
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	7021854





DataPrep Report

Approximate Distinct Count

4

chlorpropamide

categorical

Approximate Unique (%)

0.0%

Missing

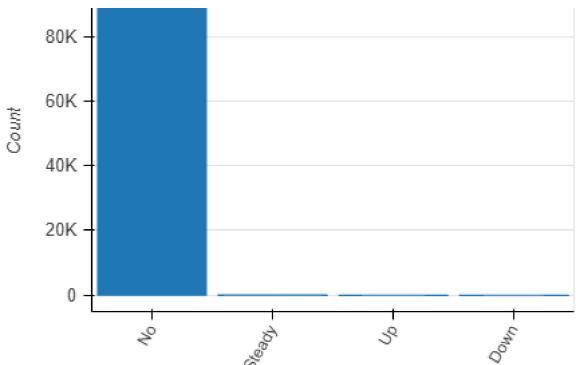
0

Missing (%)

0.0%

Memory Size

6818640

[Show Details](#)

Approximate Distinct Count

4

glimepiride

categorical

Approximate Unique (%)

0.0%

Missing

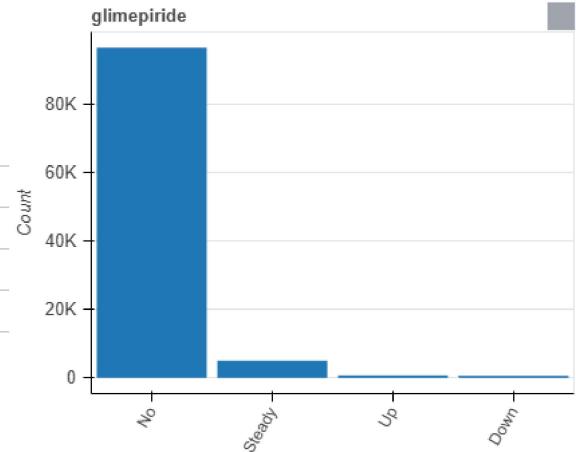
0

Missing (%)

0.0%

Memory Size

6837390

[Show Details](#)

Approximate Distinct Count

2

acetohexamide

categorical

Approximate Unique (%)

0.0%

Missing

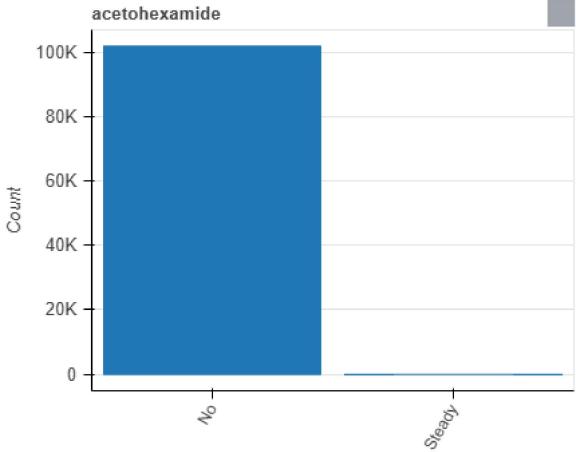
0

Missing (%)

0.0%

Memory Size

6818326

[Show Details](#)

Approximate Distinct Count

4

glipizide

categorical

Approximate Unique (%)

0.0%

Missing

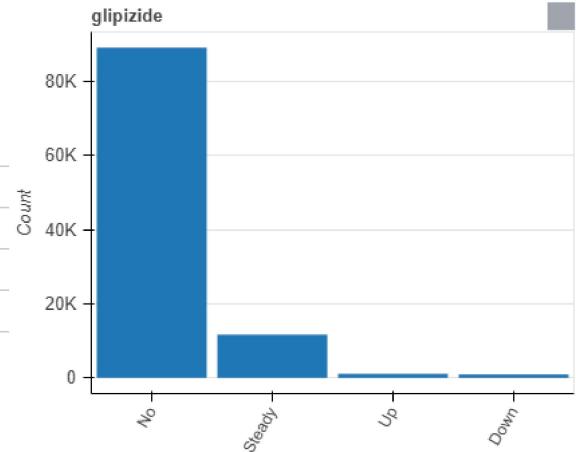
0

Missing (%)

0.0%

Memory Size

6864866

[Show Details](#)

Approximate Distinct Count

4

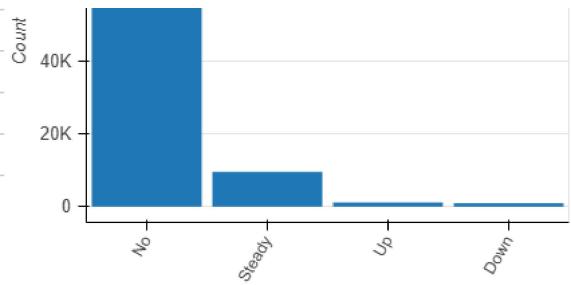
glyburide

DataPrep Report

glyburide
categorical

[Show Details](#)

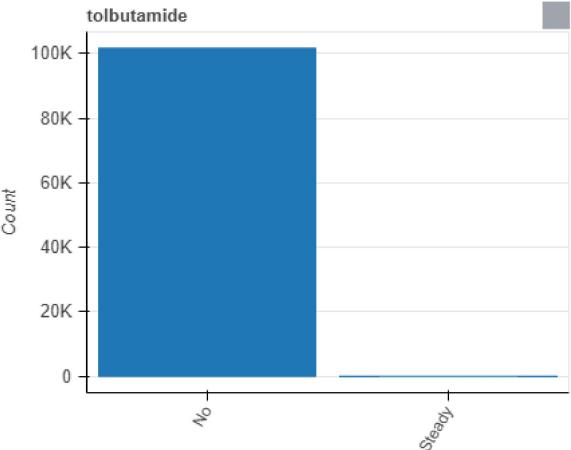
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6856546



tolbutamide
categorical

[Show Details](#)

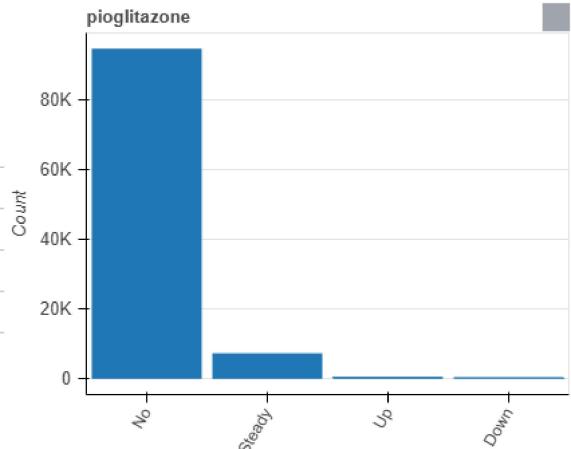
Approximate Distinct Count	2
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6818414



pioglitazone
categorical

[Show Details](#)

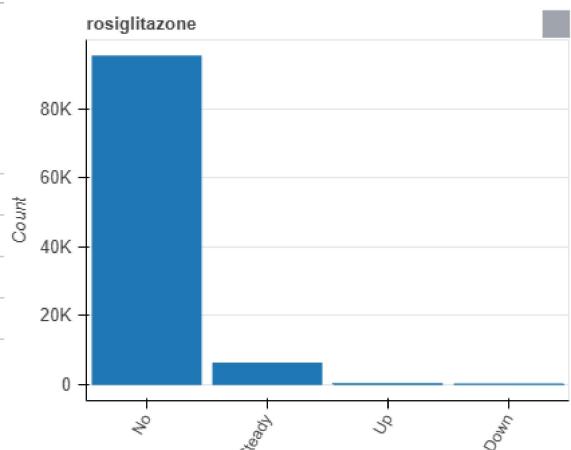
Approximate Distinct Count	4
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6846462



rosiglitazone
categorical

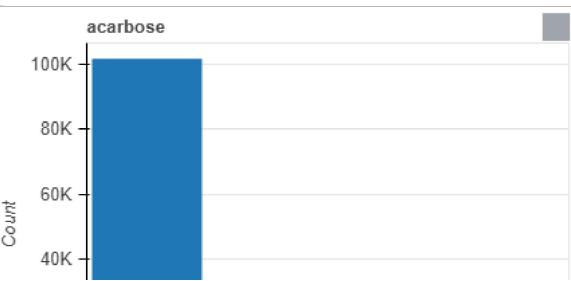
[Show Details](#)

Approximate Distinct Count	4
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6842896



acarbose
categorical

Approximate Distinct Count	4
Approximate Unique (%)	0.0%
Missing	0

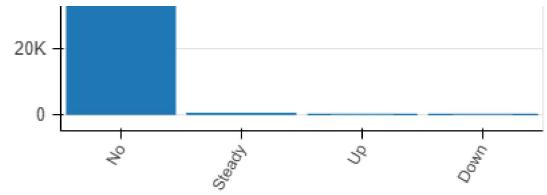


DataPrep Report

[Show Details](#)

Missing (%) 0.0%

Memory Size 6819508



Approximate Distinct Count 4

Approximate Unique (%) 0.0%

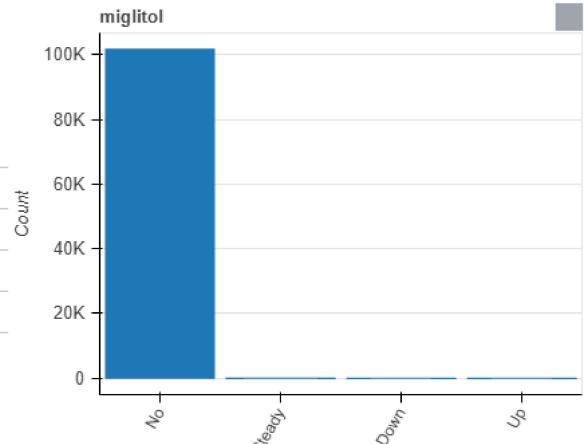
Missing 0

Missing (%) 0.0%

Memory Size 6818456

miglitol

categorical

[Show Details](#)

Approximate Distinct Count 2

Approximate Unique (%) 0.0%

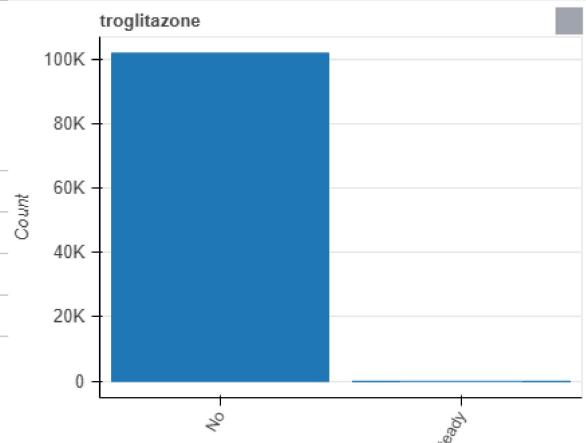
Missing 0

Missing (%) 0.0%

Memory Size 6818334

troglitazone

categorical

[Show Details](#)

Approximate Distinct Count 3

Approximate Unique (%) 0.0%

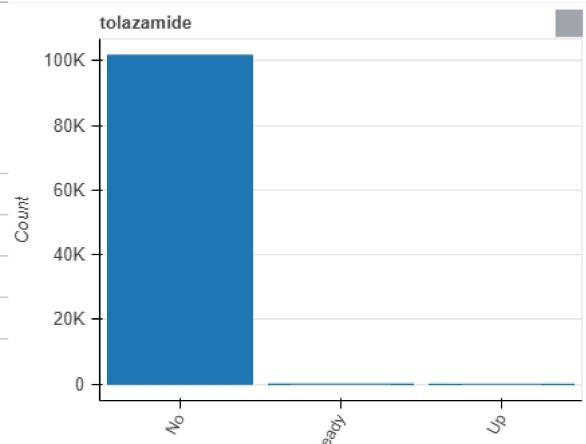
Missing 0

Missing (%) 0.0%

Memory Size 6818474

tolazamide

categorical

[Show Details](#)

Approximate Distinct Count 1

Approximate Unique (%) 0.0%

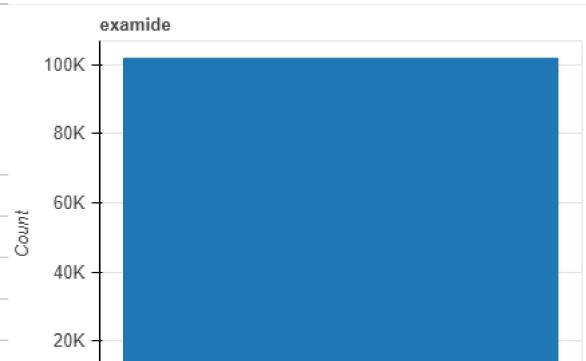
Missing 0

Missing (%) 0.0%

Memory Size 6818322

examide

categorical

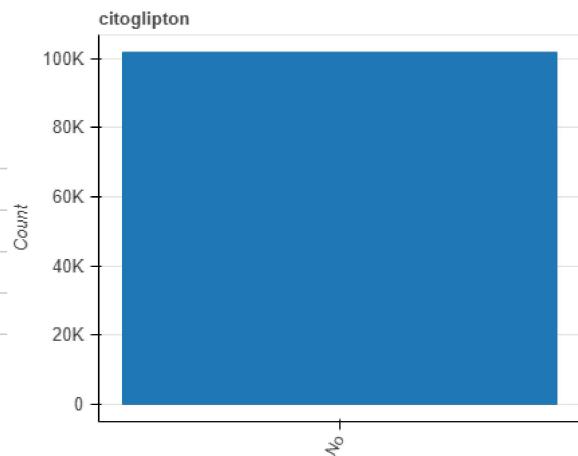
[Show Details](#)



citoglipton
categorical

Show Details

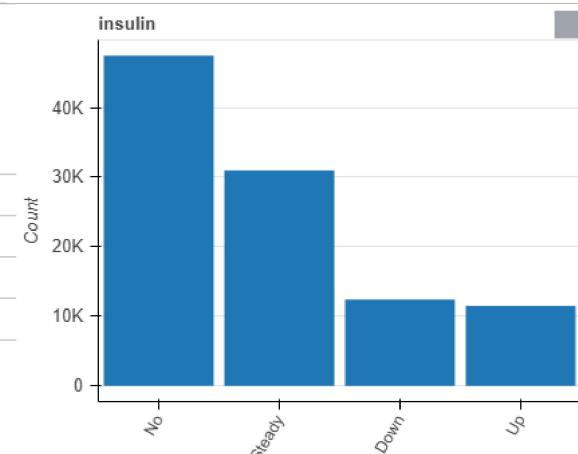
Approximate Distinct Count	1
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6818322



insulin
categorical

Show Details

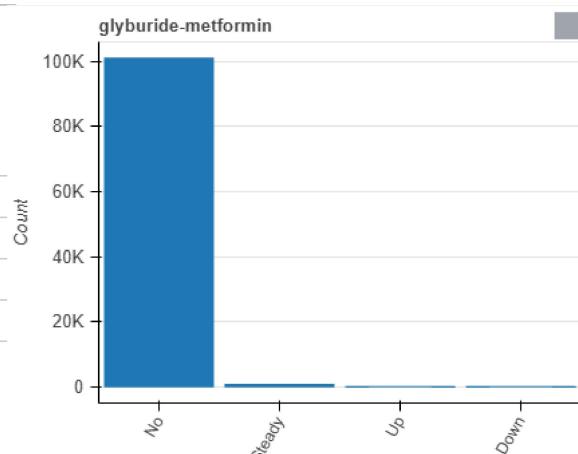
Approximate Distinct Count	4
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6966154



glyburide-metformin
categorical

Show Details

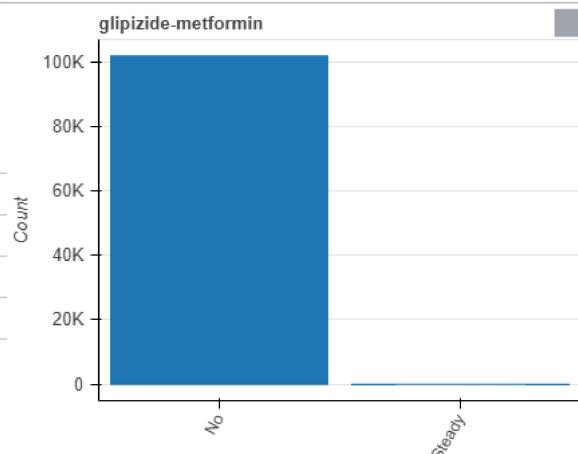
Approximate Distinct Count	4
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6821102



glipizide-metformin
categorical

Show Details

Approximate Distinct Count	2
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6818374

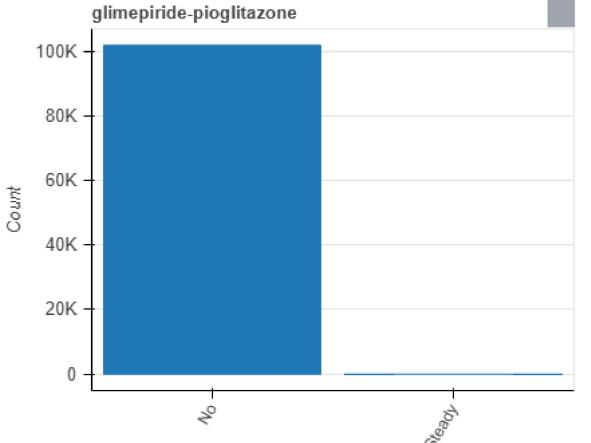


glimepiride-pioglitazone

categorical

[Show Details](#)

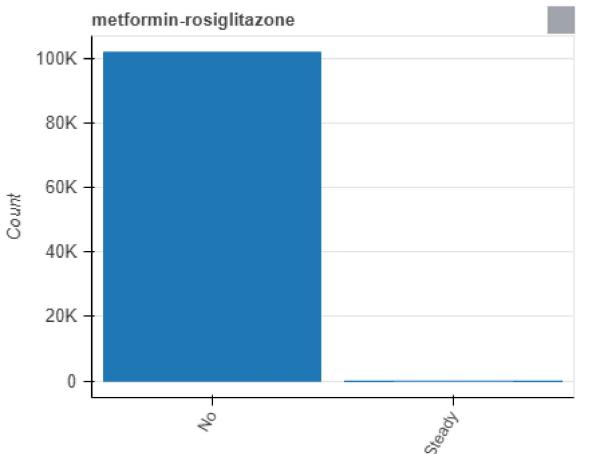
Approximate Distinct Count	2
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6818326

**metformin-rosiglitazone**

categorical

[Show Details](#)

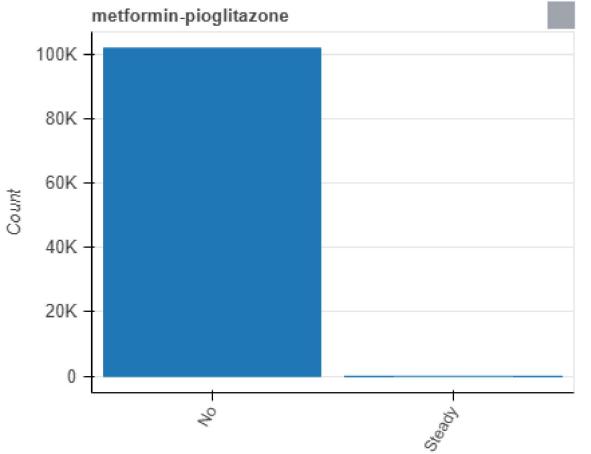
Approximate Distinct Count	2
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6818330

**metformin-pioglitazone**

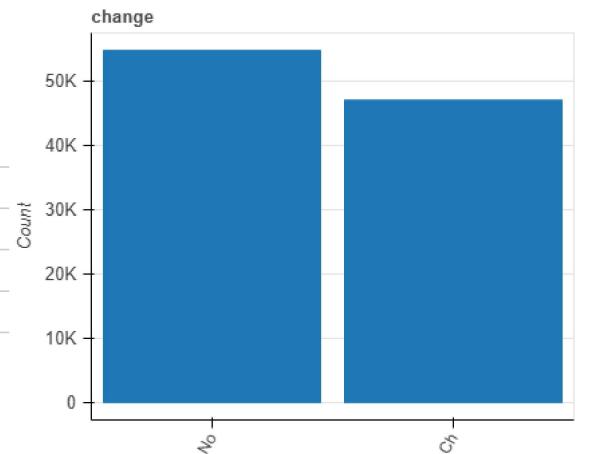
categorical

[Show Details](#)

Approximate Distinct Count	2
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6818326

**change**
categorical[Show Details](#)

Approximate Distinct Count	2
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6818322

**diabetesMed**

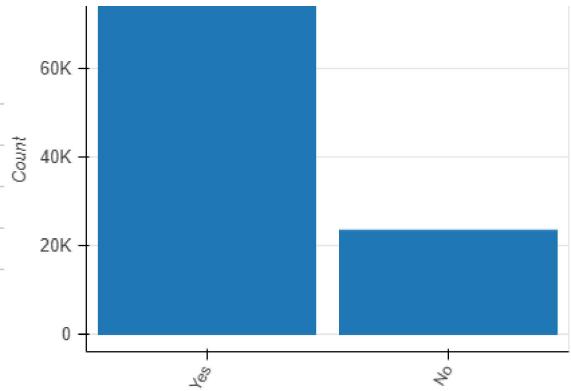
DataPrep Report

diabetesMed

categorical

[Show Details](#)

Approximate Distinct Count	2
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6896685

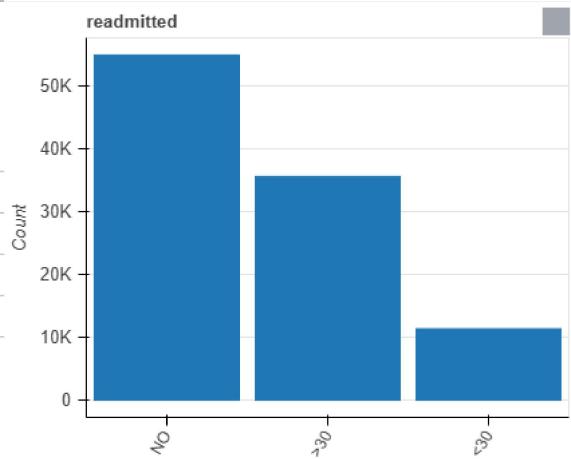


readmitted

categorical

[Show Details](#)

Approximate Distinct Count	3
Approximate Unique (%)	0.0%
Missing	0
Missing (%)	0.0%
Memory Size	6865224

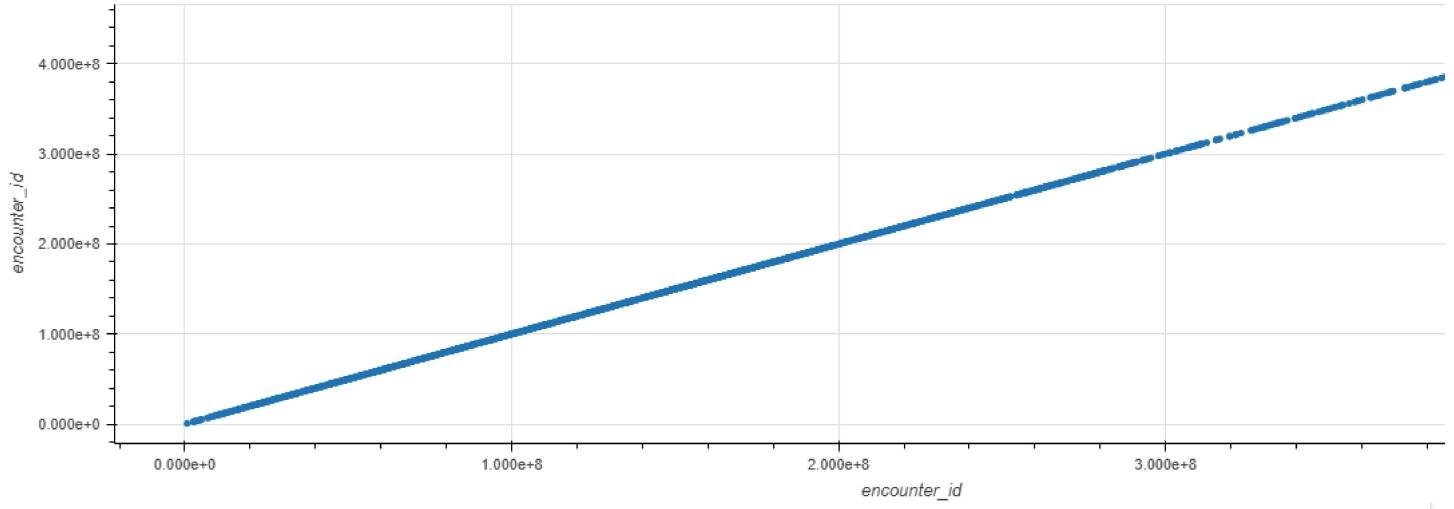


Interactions

X-Axis
encounter_id

Y-Axis
encounter_id

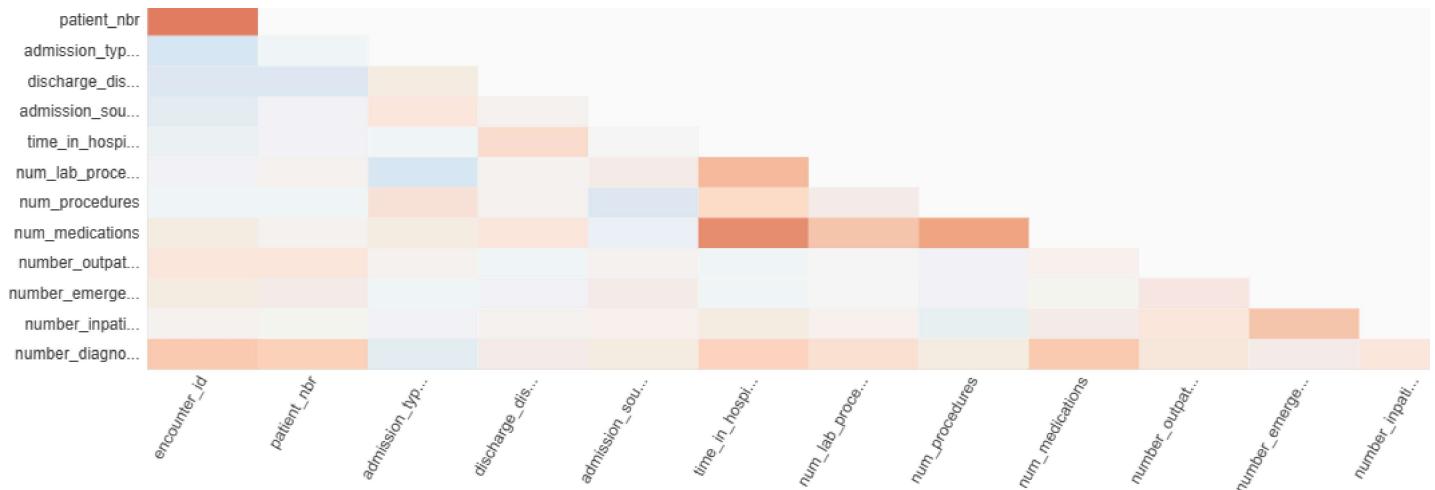
Scatter Plot



Correlations

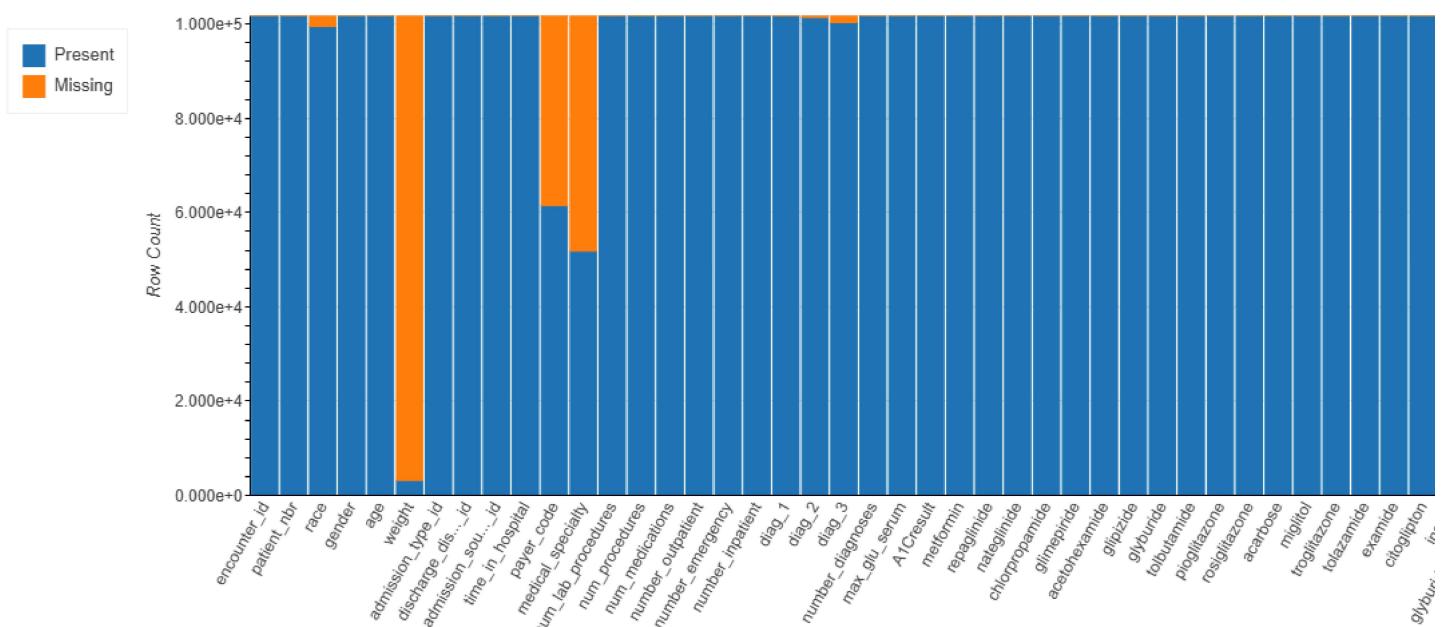
[Pearson](#)[Spearman](#)[KendallTau](#)

encounter_id



Missing Values

[Bar Chart](#) [Spectrum](#) [Heat Map](#) [Dendrogram](#)



Report generated with [DataPrep](#)