

Overview

Dataset Statistics

Number of Variables	12
Number of Rows	891
Missing Cells	866
Missing Cells (%)	8.1%
Duplicate Rows	0
Duplicate Rows (%)	0.0%
Total Size in Memory	315.7 KB
Average Row Size in Memory	362.8 B
Variable Types	Numerical: 3 Categorical: 9

Dataset Insights

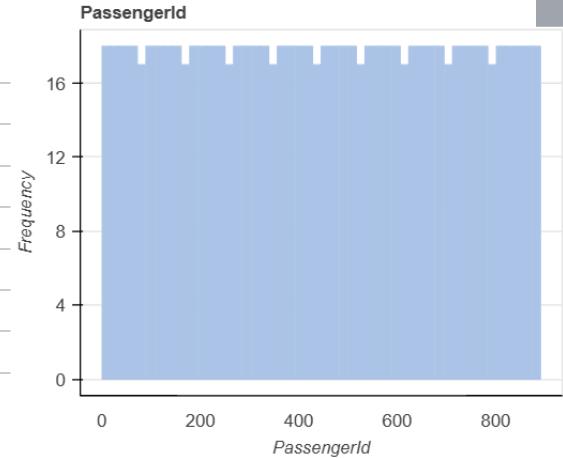
<code>PassengerId</code> is uniformly distributed	Uniform
<code>Age</code> has 177 (19.87%) missing values	Missing
<code>Cabin</code> has 687 (77.1%) missing values	Missing
<code>Fare</code> is skewed	Skewed
<code>Name</code> has a high cardinality: 891 distinct values	High Cardinality
<code>Ticket</code> has a high cardinality: 681 distinct values	High Cardinality
<code>Cabin</code> has a high cardinality: 147 distinct values	High Cardinality
<code>Survived</code> has constant length 1	Constant Length
<code>Pclass</code> has constant length 1	Constant Length
<code>SibSp</code> has constant length 1	Constant Length

1 2

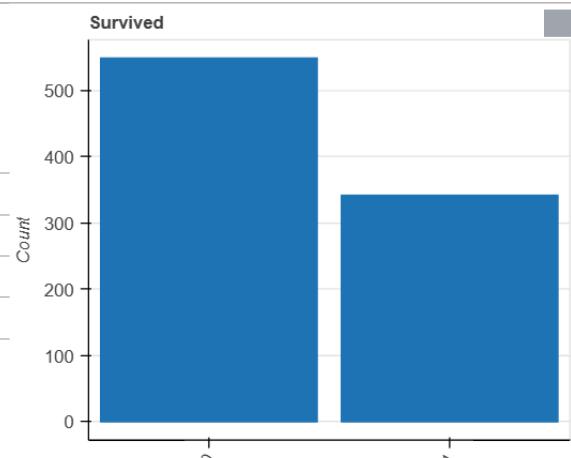
Variables

Sort by Feature order ▾ Reverse order

PassengerId numerical	Approximate Distinct Count	891
	Approximate Unique (%)	100.0%
	Missing	0
	Missing (%)	0.0%
	Infinite	0
	Infinite (%)	0.0%
	Memory Size	14256

[Show Details](#)[Show Details](#)

Approximate Distinct Count	2
Approximate Unique (%)	0.2%
Missing	0
Missing (%)	0.0%
Memory Size	58806



DataPrep Report

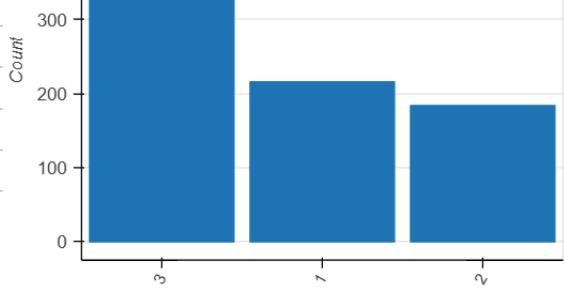
Pclass
categorical

[Show Details](#)**Approximate Distinct Count**

3

Approximate Unique (%)

0.3%



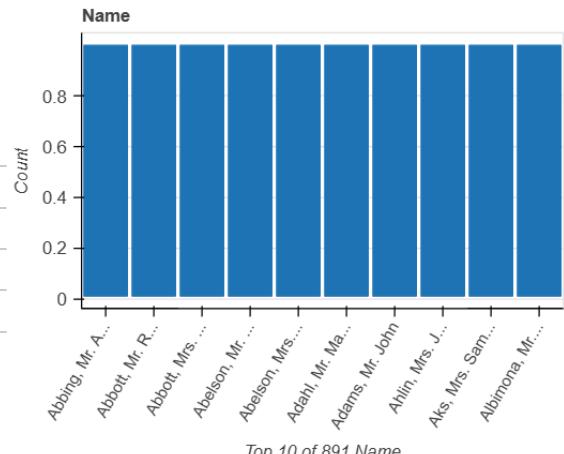
Name
categorical

[Show Details](#)**Approximate Distinct Count**

891

Approximate Unique (%)

100.0%



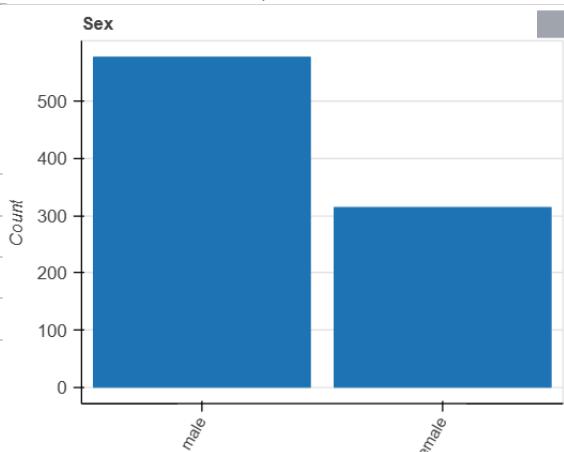
Sex
categorical

[Show Details](#)**Approximate Distinct Count**

2

Approximate Unique (%)

0.2%



Age
numerical

[Show Details](#)**Approximate Distinct Count**

88

Approximate Unique (%)

12.3%

Mean

29.6991

Missing

177

Minimum

0.42

Missing (%)

19.9%

Maximum

80

Infinite

0

Zeros

0

Infinite (%)

0.0%

Zeros (%)

0.0%

Memory Size

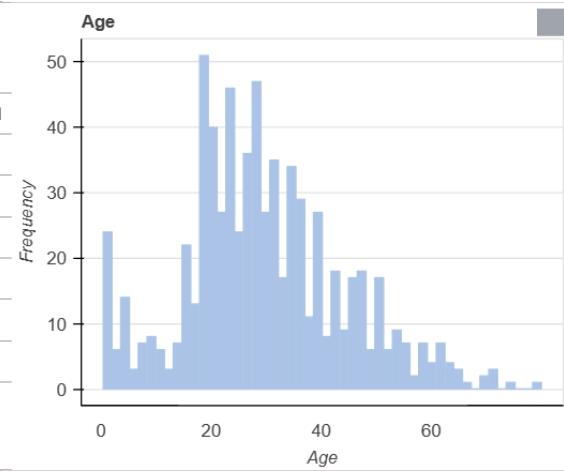
11424

Negatives

0

Negatives (%)

0.0%



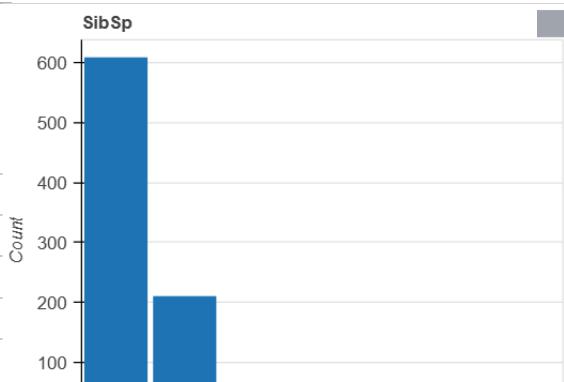
SibSp
categorical

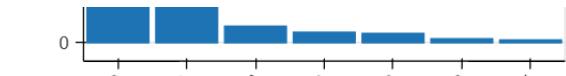
[Show Details](#)**Approximate Distinct Count**

7

Approximate Unique (%)

0.8%





Approximate Distinct Count

7

Approximate Unique (%)

0.8%

Parch

categorical

[Show Details](#)

Missing

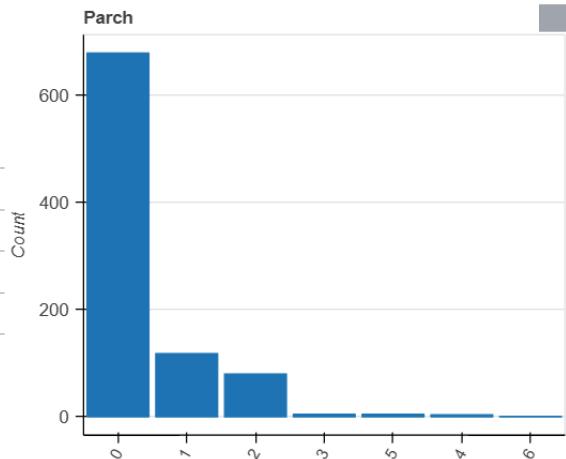
0

Missing (%)

0.0%

Memory Size

58806



Approximate Distinct Count

681

Approximate Unique (%)

76.4%

Ticket

categorical

[Show Details](#)

Missing

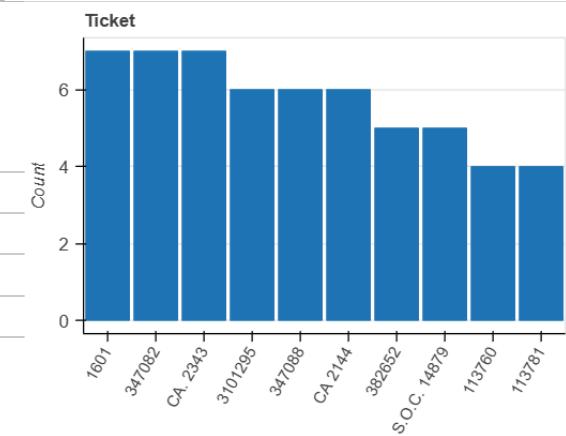
0

Missing (%)

0.0%

Memory Size

63930



Approximate Distinct Count

248

Mean

32.2042

Approximate Unique (%)

27.8%

Fare

numerical

[Show Details](#)

Missing

0

Missing (%)

0.0%

Infinite

0

Infinite (%)

0.0%

Memory Size

14256

Minimum

0

Maximum

512.3292

Zeros

15

Zeros (%)

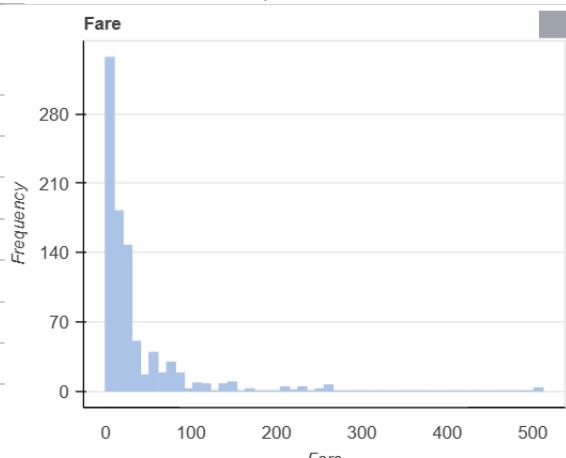
1.7%

Negatives

0

Negatives (%)

0.0%



Approximate Distinct Count

147

Approximate Unique (%)

72.1%

Cabin

categorical

[Show Details](#)

Missing

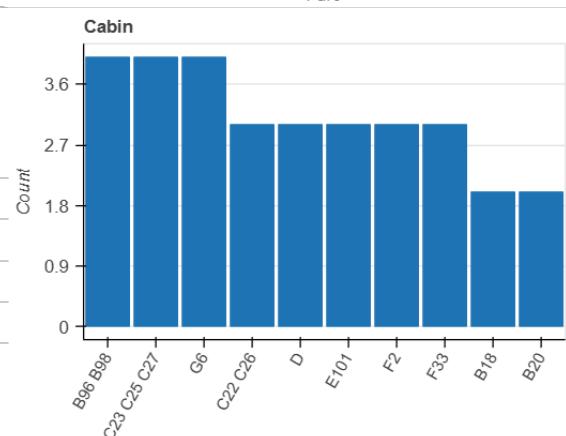
687

Missing (%)

77.1%

Memory Size

13992



DataPrep Report

Embarked
 categorical
[Show Details](#)

Approximate Distinct Count

3

Approximate Unique (%)

0.3%

Missing

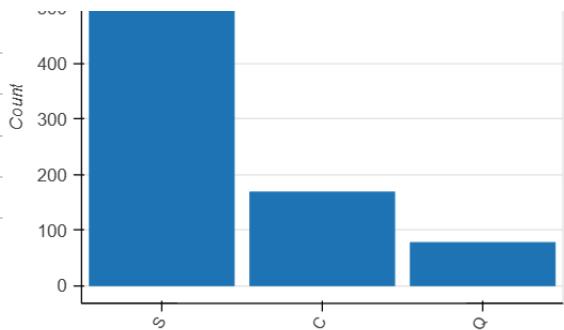
2

Missing (%)

0.2%

Memory Size

58674



Interactions

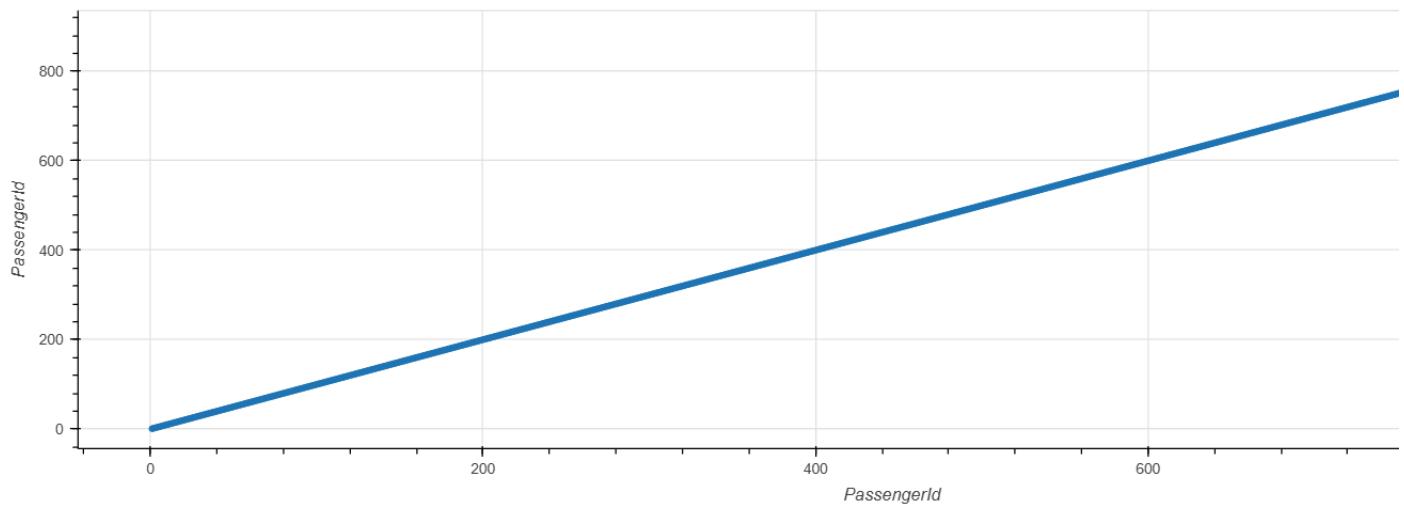
X-Axis

PassengerId

Y-Axis

PassengerId

Scatter Plot



Correlations

Pearson

Spearman

KendallTau

PassengerId

Survived

Pclass

Age

SibSp

Parch

Fare

PassengerId

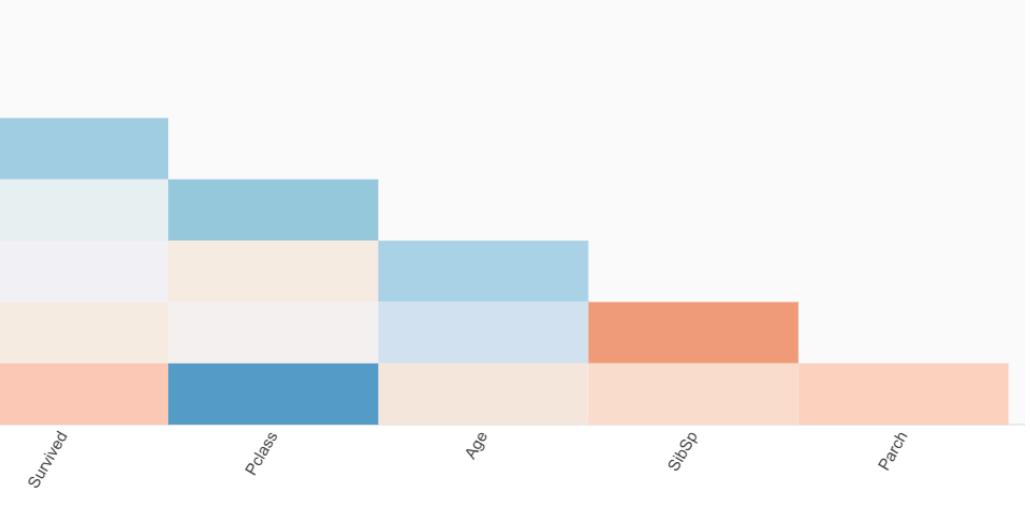
Survived

Pclass

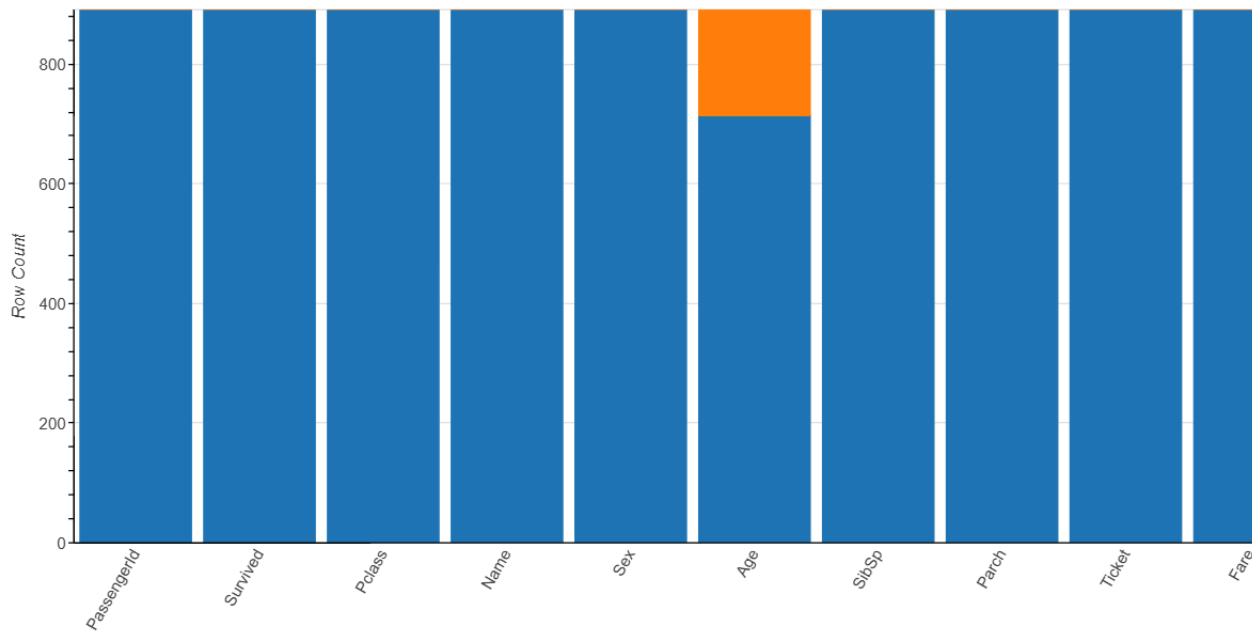
Age

SibSp

Parch



Missing Values

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

Report generated with [DataPrep](#)