Exploratory Data Analysis (EDA)

Filter-based, on-demand exploratory analysis

Dataset Overview

Rows
891

Columns
14

Column Data Types

	Column	Data Type
PassengerId	PassengerId	int64
Survived	Survived	int64
Pclass	Pclass	int64
Name	Name	object
Sex	Sex	int64
Age	Age	category
SibSp	SibSp	int64
Parch	Parch	int64
Ticket	Ticket	object
Fare	Fare	float64

Feature Type Count


```
{
  "numeric_features" : 7
  "categorical_features" : 7
  "numeric_columns" : [
    0 : "PassengerId"
    1 : "Survived"
    2 : "Pclass"
    3 : "Sex"
    4 : "SibSp"
    5 : "Parch"
    6 : "Fare"
  ]
  "categorical_columns" : [
    0 : "Name"
    1 : "Age"
    2 : "Ticket"
    3 : "Cabin"
    4 : "Embarked_C"
    5 : "Embarked_Q"
    6 : "Embarked_S"
  ]
}
```


Memory Usage: 0.21 MB

	Column	Memory (KB)
0	Index	0.13
1	PassengerId	6.96
2	Survived	6.96
3	Pclass	6.96
4	Name	66.1
5	Sex	6.96
6	Age	1.13
7	SibSp	6.96
8	Parch	6.96
9	Ticket	48.51

Constant Columns


```
[]
```


High Cardinality Columns

	Column	Unique Values
0	Name	891
1	Ticket	681
2	Cabin	147

Descriptive Statistics

Select numeric columns

Survived × Fare ×

	Mean	Median	Min	Max	Range	Variance	Std Dev	Q1	Q2 (Median)	Q3	IQR
Survived	0.3838	0	0	1	1	0.2368	0.4866	0	0	1	1
Fare	0	-0.3574	-0.6484	9.6672	10.3156	1.0011	1.0006	-0.4891	-0.3574	-0.0242	0.4649

Categorical Data Analysis

Select categorical column

Age

	Age	Count
0	Adult	526
1	Young	301
2	Senior	64

Dominance:


```
{
  "Most Frequent Category" : "Adult"
  "Dominance Ratio" : 0.5903
}
```


	Category	Percentage
	empty	
