Data Analysis Report

Generated on: 2025-10-18 11:42

Dataset Overview

Total Rows	891
Total Columns	26
Memory Usage	0.39 MB
Missing Values	13338

Column Summary

Column Name	Data Type	Missing Values	Unique Values
PassengerId	int64	0	891
Survived	int64	0	2
Pclass	int64	0	3
Name	object	0	891
Sex	object	0	2
Age	float64	177	88
SibSp	int64	0	7
Parch	int64	0	7
Ticket	object	0	681
Fare	float64	0	248
Cabin	object	687	147
Embarked	object	2	3
Unnamed: 12	float64	891	0
Unnamed: 13	float64	891	0
Unnamed: 14	float64	891	0
Unnamed: 15	float64	891	0
Unnamed: 16	float64	891	0
Unnamed: 17	float64	891	0
Unnamed: 18	float64	891	0
Unnamed: 19	float64	891	0

Unnamed: 20	float64	891	0
Unnamed: 21	float64	891	0
Unnamed: 22	float64	891	0
Unnamed: 23	float64	891	0
Unnamed: 24	float64	891	0
Unnamed: 25	object	889	2

Basic Statistics

Statistics for PassengerId

Metric	Value
Count	891
Mean	446.00
Std Dev	257.35
Min	1.00
25%	223.50
50%	446.00
75%	668.50
Max	891.00

Statistics for Survived

Metric	Value
Count	891
Mean	0.38
Std Dev	0.49
Min	0.00
25%	0.00
50%	0.00
75%	1.00
Max	1.00

Statistics for Pclass

Metric	Value
Count	891

Mean	2.31
Std Dev	0.84
Min	1.00
25%	2.00
50%	3.00
75%	3.00
Max	3.00

Key Insights

- Dataset contains 891 rows and 26 columns
- Dataset has 13338 missing values that may need handling
- Found 20 numeric columns for analysis
- Found 6 categorical columns
- 'male' is the most frequent category in Sex (577 occurrences)