Exploratory Data Analysis (EDA) Report - Synthetic Titanic Dataset

Dataset Info

RangeIndex: 891 entries, 0 to 890
Data columns (total 9 columns):
Column Non-Null Count Dtype
--- ------ -------------- -----
0 survived 891 non-null int64
1 pclass 891 non-null int64
2 sex 891 non-null object
3 age 713 non-null float64
4 sibsp 891 non-null int64
5 parch 891 non-null int64
6 fare 891 non-null float64
7 embarked 891 non-null object
8 class 891 non-null object
dtypes: float64(2), int64(4), object(3)
memory usage: 62.8+ KB

Summary Statistics

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class
count	891.000000	891.000000	891	713.000000	891.000000	891.000000	891.000000	891	891
unique	NaN	NaN	2	NaN	NaN	NaN	NaN	3	3
top	NaN	NaN	male	NaN	NaN	NaN	NaN	S	Third
freq	NaN	NaN	583	NaN	NaN	NaN	NaN	684	505
mean	0.346801	2.316498	NaN	30.616715	0.498316	0.383838	47.630572	NaN	NaN
std	0.476219	0.847168	NaN	13.885098	0.689200	0.609593	35.860358	NaN	NaN
min	0.000000	1.000000	NaN	-12.273170	0.000000	0.000000	0.060000	NaN	NaN
25%	0.000000	1.500000	NaN	21.490492	0.000000	0.000000	19.070000	NaN	NaN
50%	0.000000	3.000000	NaN	30.292411	0.000000	0.000000	38.970000	NaN	NaN
75%	1.000000	3.000000	NaN	39.608458	1.000000	1.000000	69.710000	NaN	NaN
max	1.000000	3.000000	NaN	74.703506	4.000000	3.000000	186.910000	NaN	NaN

Missing Values

survived 0
pclass 0
sex 0
age 178
sibsp 0
parch 0
fare 0
embarked 0
class 0

Value Counts - Passenger Class

Third 505
First 223
Second 163

Observations

- Total records: 891
- Survival rate (approx): 34.68%
- Missing ages: 178 (19.98%)
- Class distribution (First/Second/Third counts): [223, 163, 505]
- - Fare is right-skewed (some high-value fares).
- - Age has missing values (about 20%).
- - Survival differs by sex (females tend to survive more in many Titanic datasets).

Visualizations

Age Distribution

Correlation Matrix (numeric features)

	survived	pclass	age	sibsp	parch	fare
survived	1.00	0.01	-0.06	-0.00	0.03	-0.03
pclass	0.01	1.00	0.06	-0.02	-0.08	-0.01
age	-0.06	0.06	1.00	-0.11	0.01	-0.05
sibsp	-0.00	-0.02	-0.11	1.00	-0.01	0.01
parch	0.03	-0.08	0.01	-0.01	1.00	-0.01
fare	-0.03	-0.01	-0.05	0.01	-0.01	1.00

Survival Rate by Sex