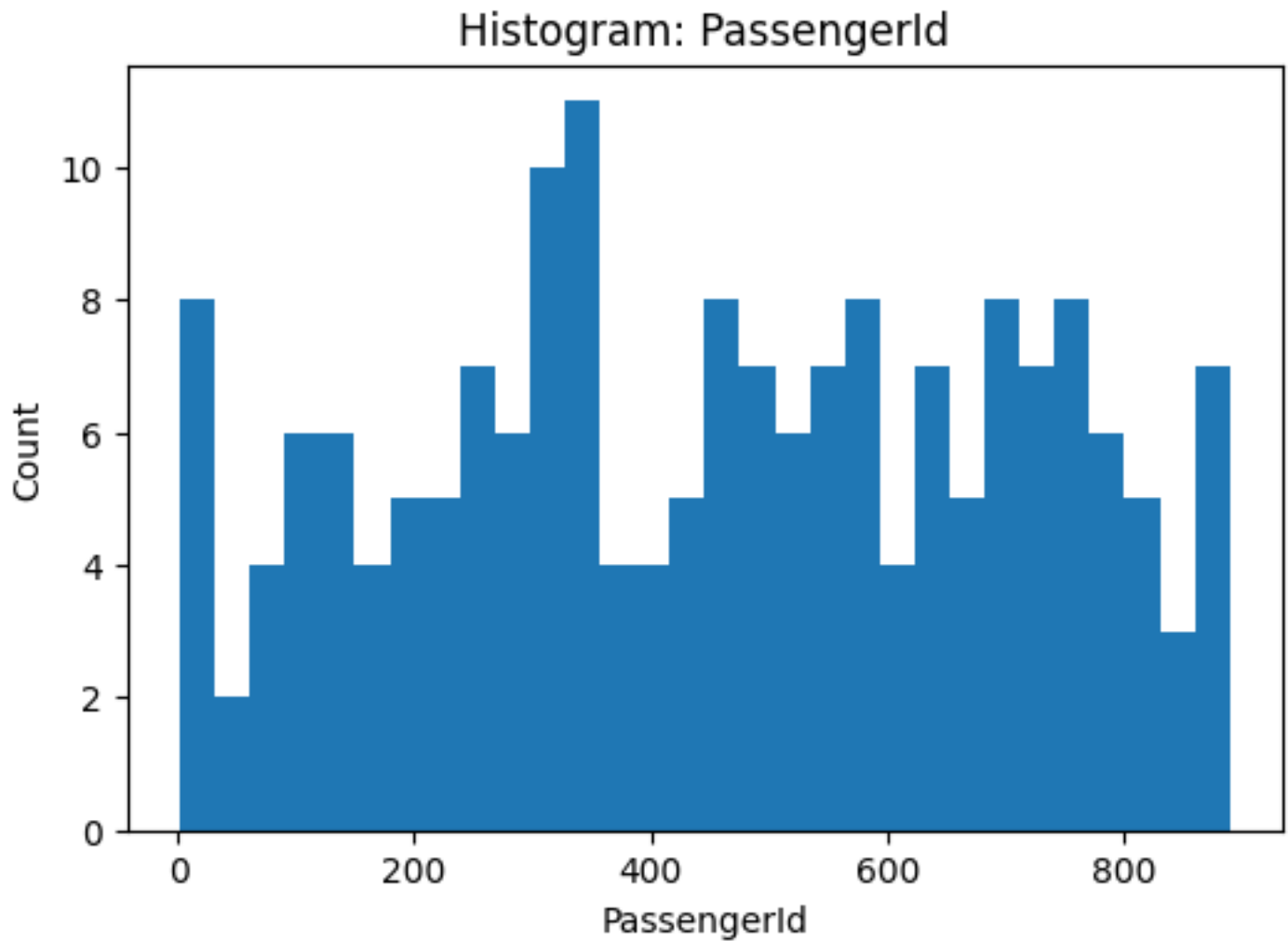
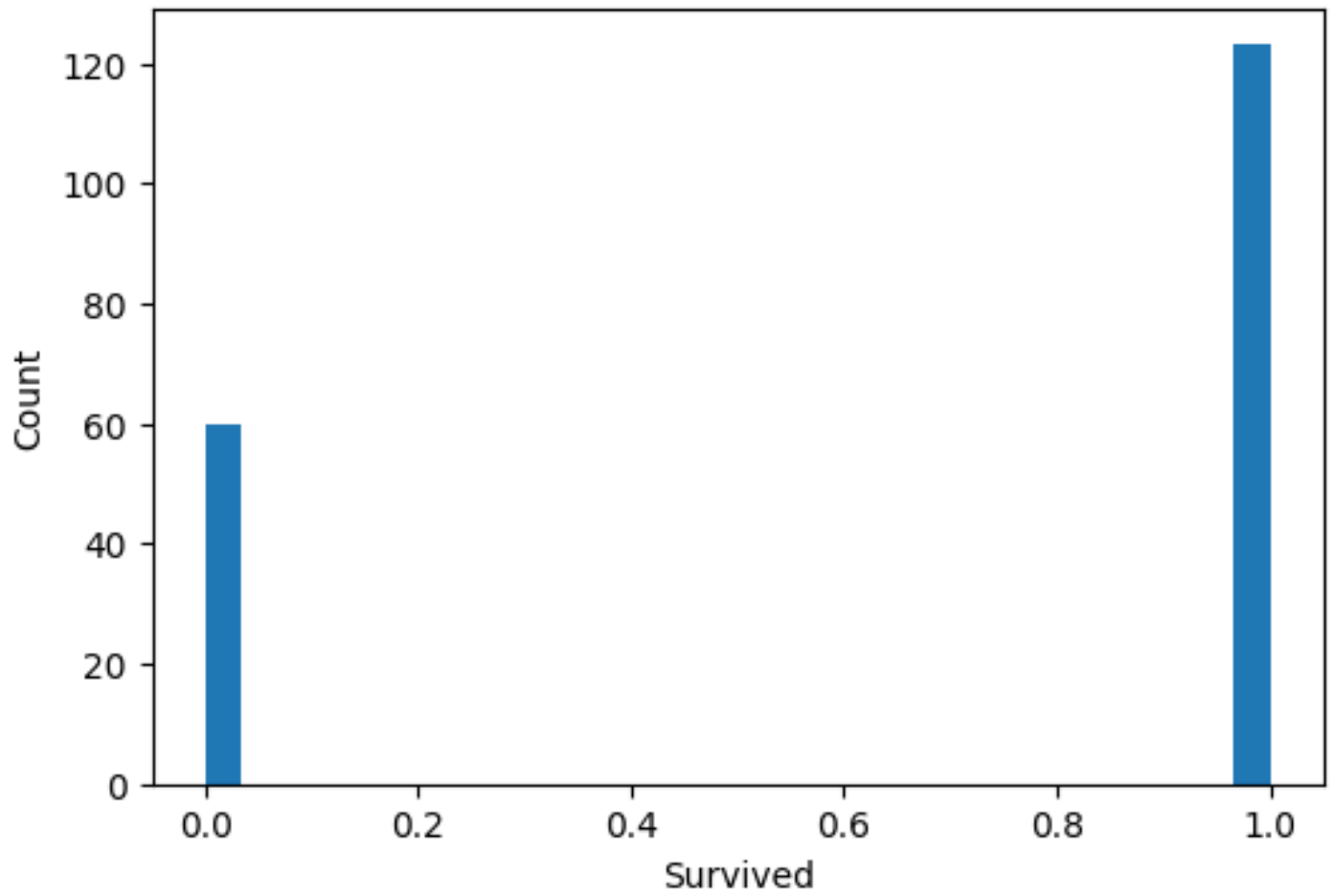


Titanic EDA Report (Final)

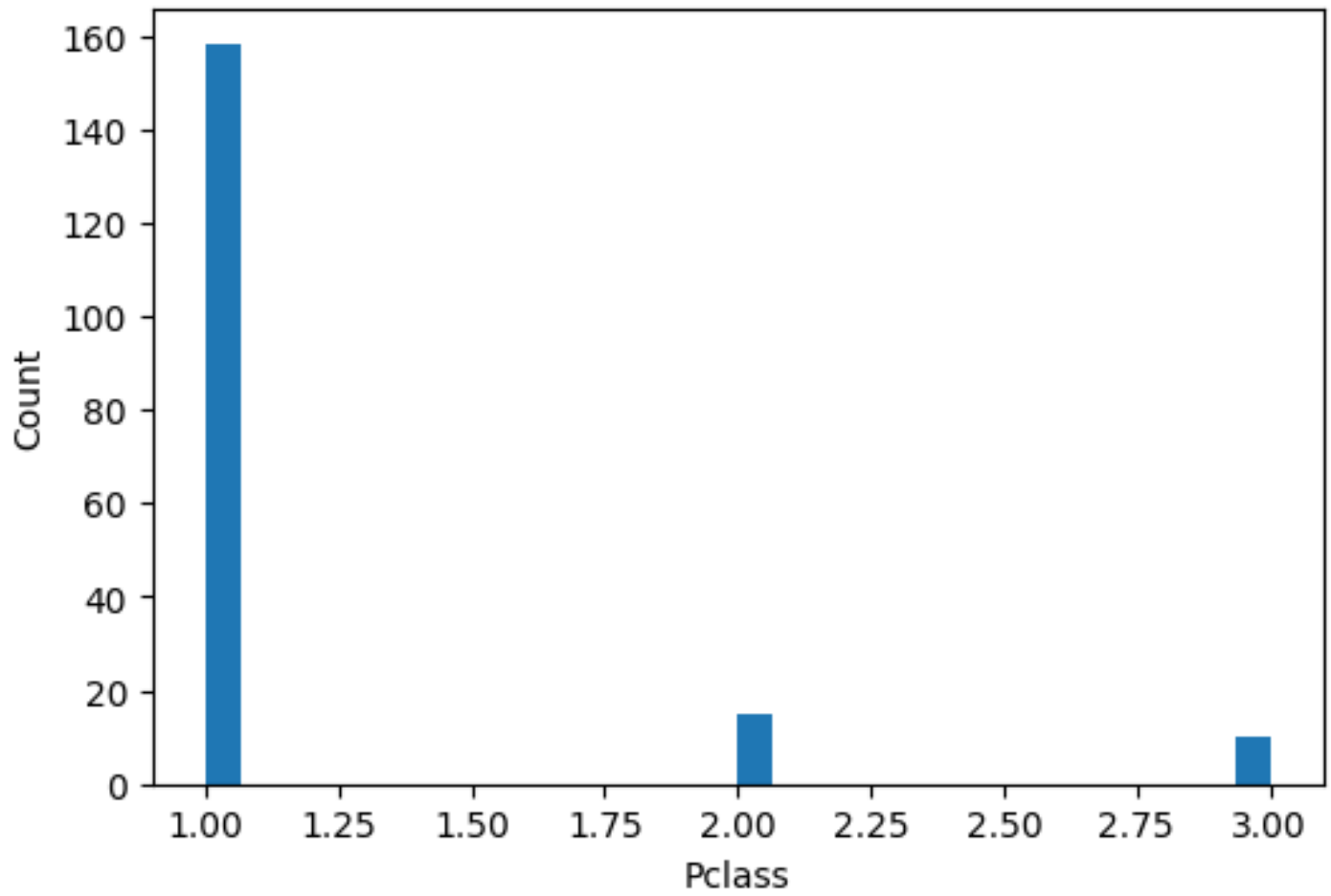
1. Cabin is almost entirely missing — not useful without heavy feature-engineering.
2. Age has ~177 missing values → needs imputation (median or model-based).
3. Sex and Pclass are strong predictors of survival:
 - Females survived at $\approx 74\%$, Males $\approx 19\%$.
 - 1st class survival $\approx 63\%$, 2nd $\approx 47\%$, 3rd $\approx 24\%$.
4. Fare and Pclass are moderately correlated ($r \approx 0.55$).
5. SibSp and Parch moderately correlated ($r \approx 0.41$).
6. Recommendations: impute Age, feature-engineer FamilySize, encode categorical variables, log-transform Fare.



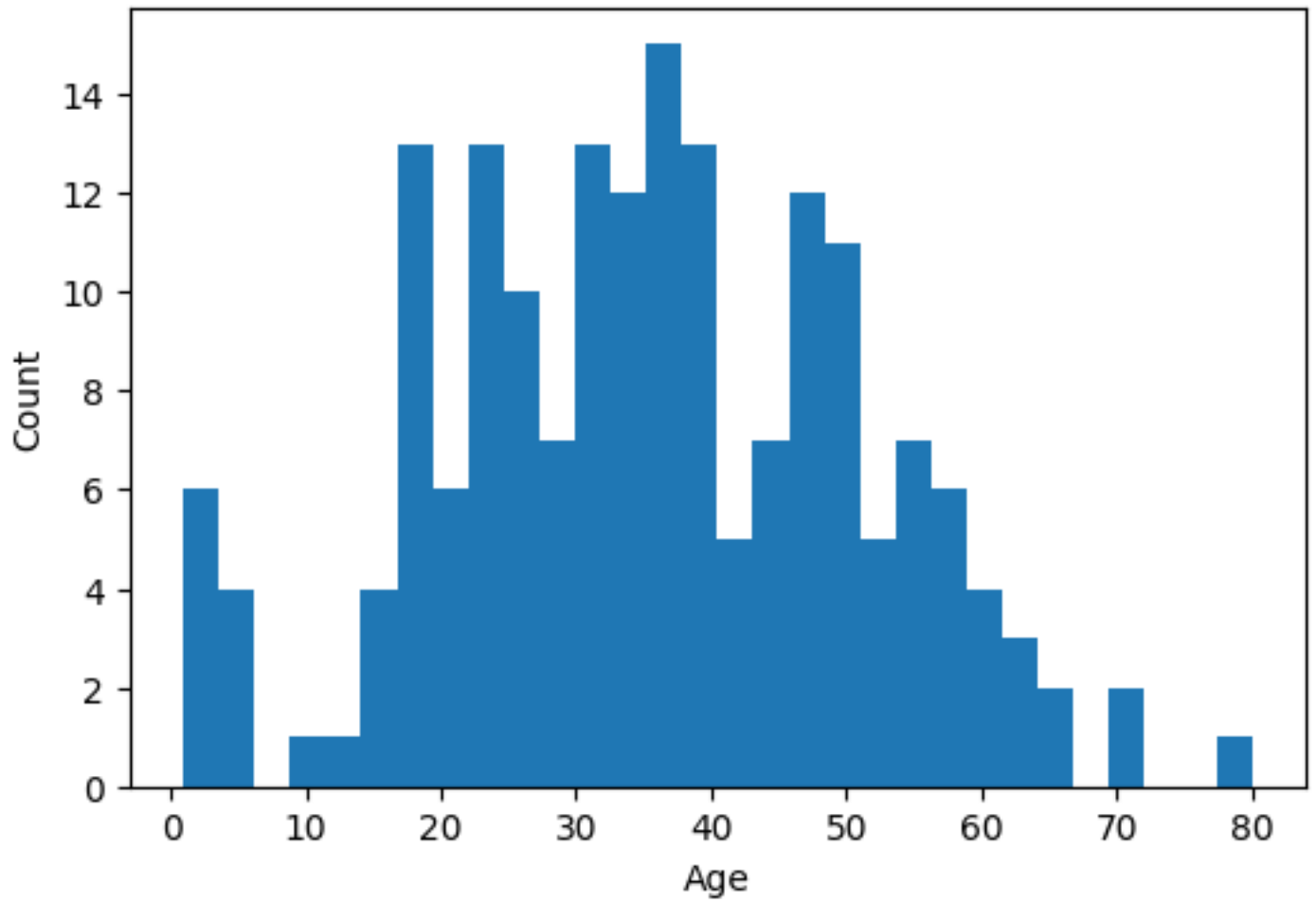
Histogram: Survived



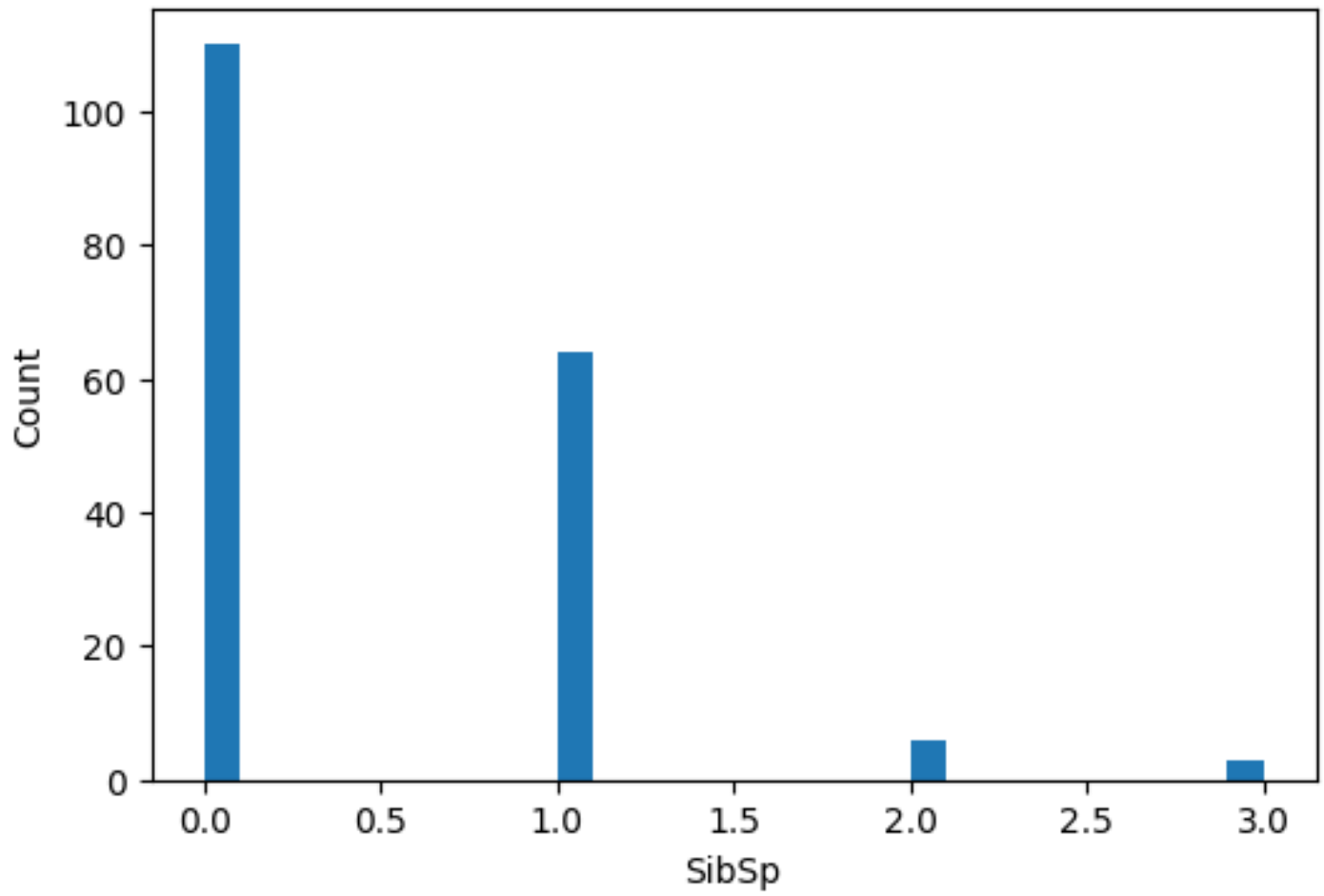
Histogram: Pclass



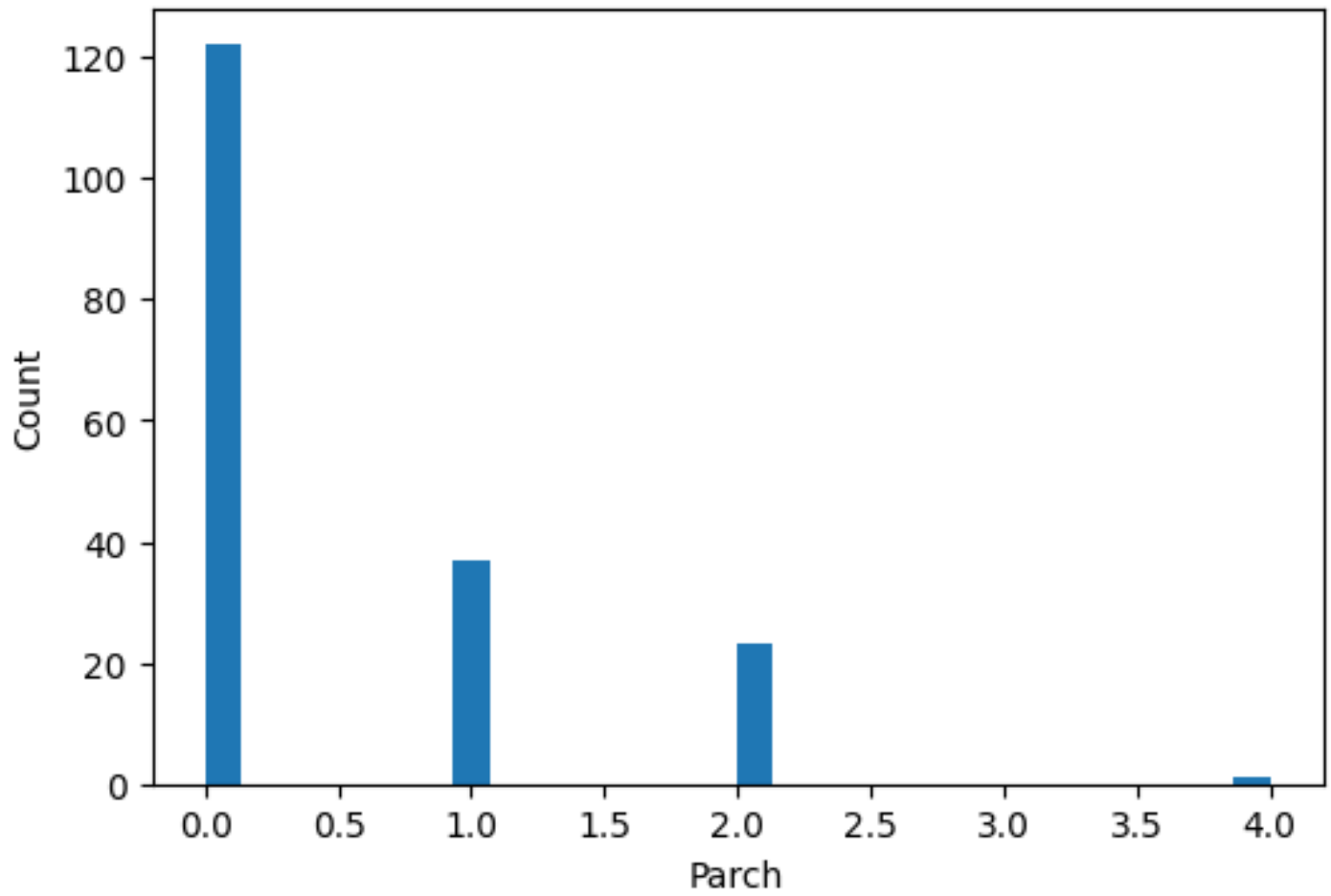
Histogram: Age



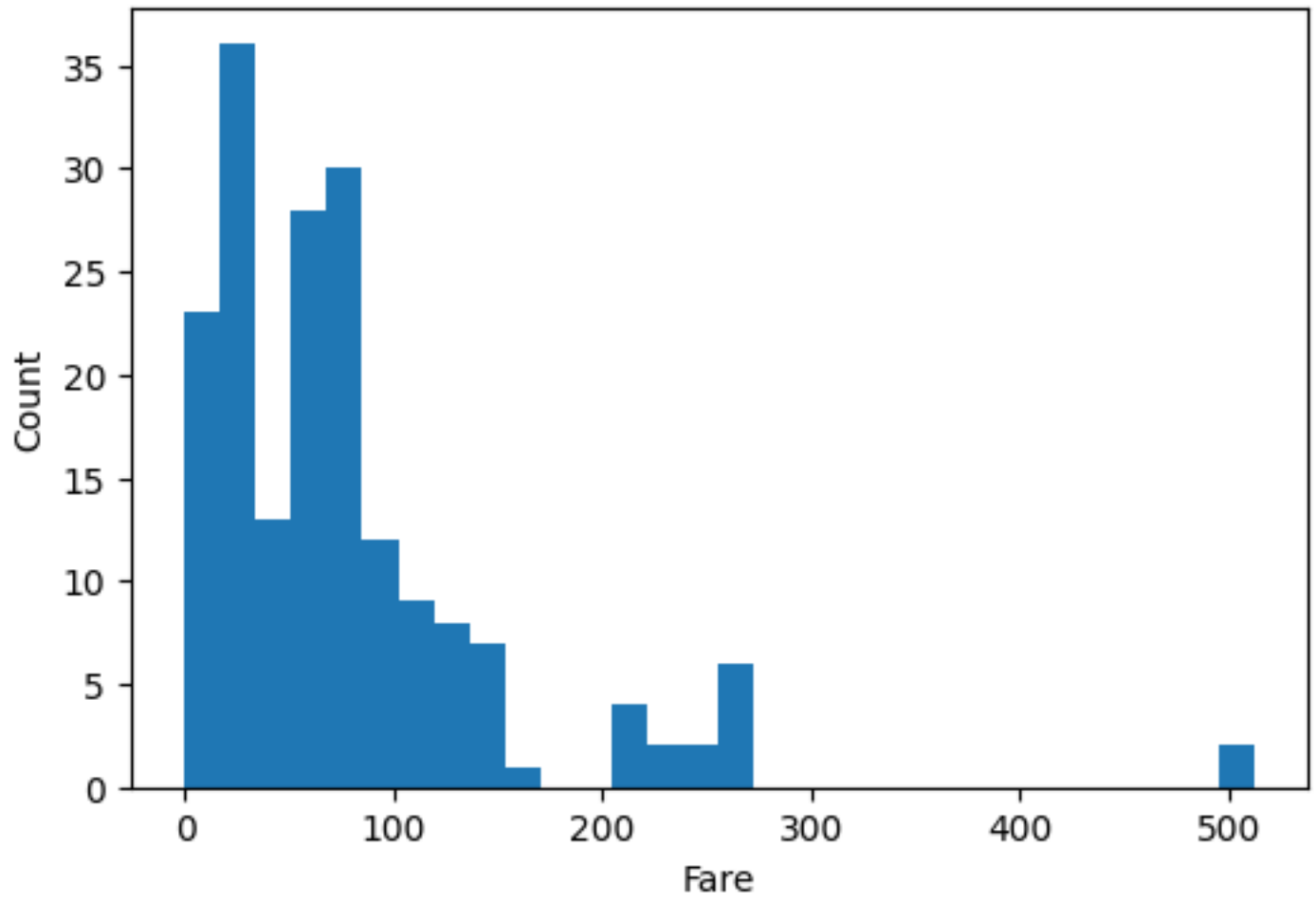
Histogram: SibSp



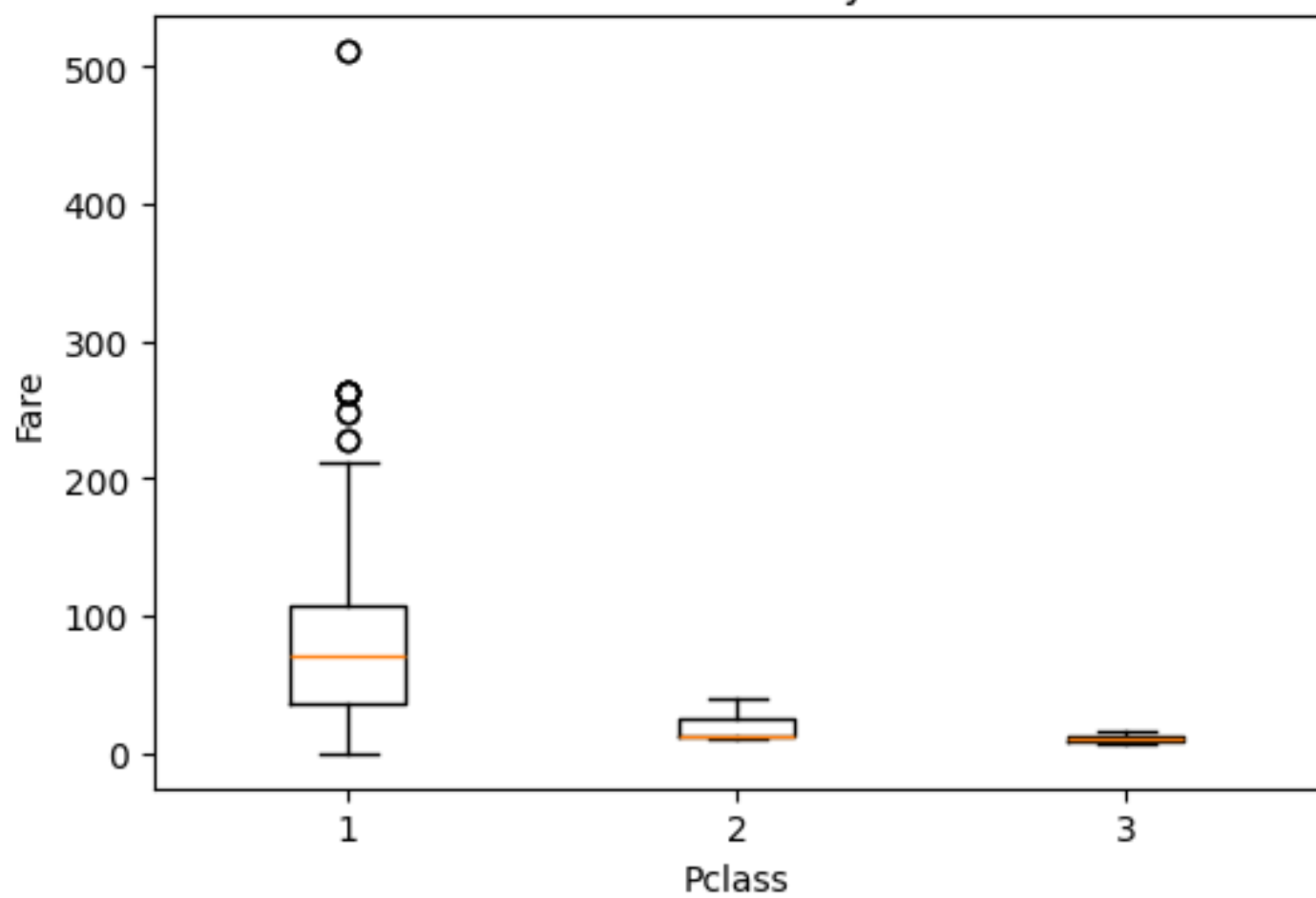
Histogram: Parch



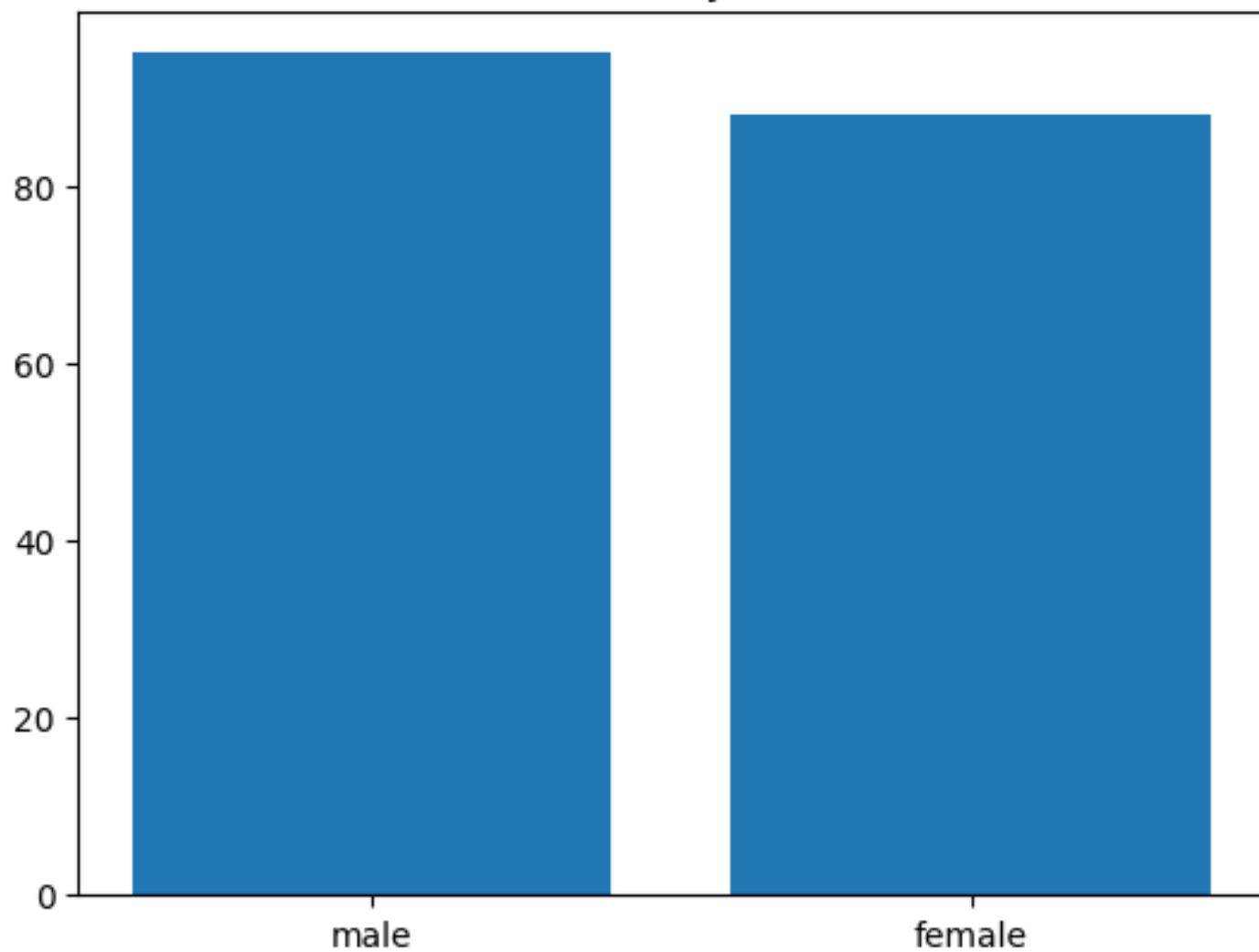
Histogram: Fare



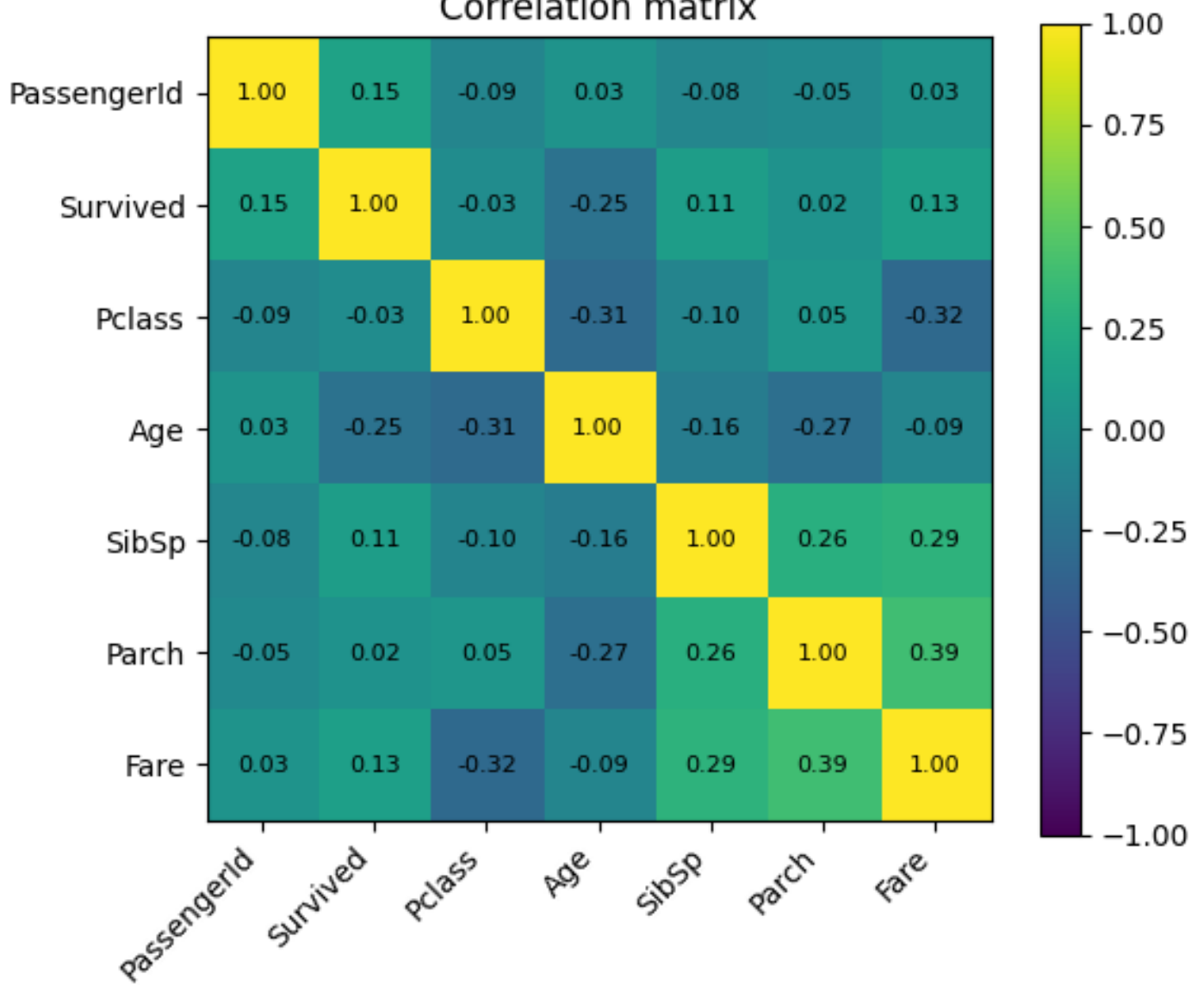
Fare distribution by Pclass



Counts by Sex



Correlation matrix



Age vs Fare by Survived

