In [2]:
```
# Importing Libraries

import pandas as pd
from scipy.stats import shapiro, normaltest, anderson, pearsonr, spearmanr, kendalltau, chi2_contingency
from statsmodels.tsa.stattools import adfuller, kpss
from scipy.stats import ttest_ind, ttest_rel, f_oneway, mannwhitneyu, wilcoxon, kruskal, friedmanchisquare
```

In [13]:
```
# Load Titanic dataset

titanic_df= pd.read_csv('C:/Users/squir/Downloads/tested.csv')
```

In [14]:
```
titanic_df.head()
```

Out[14]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	892	0	3	Kelly, Mr. James	male	34.5	0	0	330911	7.8292	NaN	Q
1	893	1	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.0000	NaN	S
2	894	0	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.6875	NaN	Q
3	895	0	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.6625	NaN	S
4	896	1	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298	12.2875	NaN	S

In [15]:
```
# Drop irrelevant columns
titanic_df = titanic_df.drop(['Name', 'Ticket', 'Cabin', 'Embarked'], axis=1)
```

In [17]:
```
# Handle missing values
titanic_df['Age'].fillna(titanic_df['Age'].median(), inplace=True)
titanic_df['Fare'].fillna(titanic_df['Fare'].median(), inplace=True)
```

In [18]:
```
# Convert categorical variables to numerical
titanic_df['Sex'] = titanic_df['Sex'].map({'male': 0, 'female': 1})
```

In [19]:
```
# Create a new variable 'Survived' to represent survival as 1 and not survival as 0
titanic_df['Survived'] = titanic_df['Survived'].astype(int)
```

In [20]:
```
# Normality Tests
# Shapiro-Wilk Test
stat, p_value = shapiro(titanic_df['Age'])
print(f"Shapiro-Wilk Test: Statistics={stat}, p-value={p_value}")

# D'Agostino's K^2 Test
stat, p_value = normaltest(titanic_df['Age'])
print(f"D'Agostino's K^2 Test: Statistics={stat}, p-value={p_value}")

# Anderson-Darling Test
result = anderson(titanic_df['Age'])
print(f"Anderson-Darling Test: Statistic={result.statistic}, Critical Values={result.critical_values}, Significance Level={result
```

```
Shapiro-Wilk Test: Statistics=0.9353150129318237, p-value=1.7022099875474428e-12
D'Agostino's K^2 Test: Statistics=34.81555255420222, p-value=2.7535871142178257e-08
Anderson-Darling Test: Statistic=12.460808507225352, Critical Values=[0.571 0.65  0.78  0.909 1.082], Significance Level=[15.
10.   5.   2.5  1. ]
```

In [21]:
```
# Correlation Tests
# Pearson's Correlation Coefficient
correlation, p_value = pearsonr(titanic_df['Age'], titanic_df['Fare'])
print(f"Pearson's Correlation Coefficient: Correlation={correlation}, p-value={p_value}")

# Spearman's Rank Correlation
correlation, p_value = spearmanr(titanic_df['Age'], titanic_df['Fare'])
print(f"Spearman's Rank Correlation: Correlation={correlation}, p-value={p_value}")

# Kendall's Rank Correlation
correlation, p_value = kendalltau(titanic_df['Age'], titanic_df['Fare'])
print(f"Kendall's Rank Correlation: Correlation={correlation}, p-value={p_value}")

# Chi-Squared Test
contingency_table = pd.crosstab(titanic_df['Survived'], titanic_df['Sex'])
chi2_stat, p_value, dof, expected = chi2_contingency(contingency_table)
print(f"Chi-Squared Test: Chi2 Statistic={chi2_stat}, p-value={p_value}, Degrees of Freedom={dof}")
```

```
Pearson's Correlation Coefficient: Correlation=0.34235685018571027, p-value=6.147154025484477e-13
Spearman's Rank Correlation: Correlation=0.27724790736124283, p-value=8.177127214177605e-09
Kendall's Rank Correlation: Correlation=0.18843374022157644, p-value=2.7225756670044467e-08
Chi-Squared Test: Chi2 Statistic=413.6897405343716, p-value=5.767311139789629e-92, Degrees of Freedom=1
```

In [22]:
```
# Stationary Tests
# Augmented Dickey-Fuller
result = adfuller(titanic_df['Fare'])
print(f"Augmented Dickey-Fuller Test: ADF Statistic={result[0]}, p-value={result[1]}, Critical Values={result[4]}")

# Kwiatkowski-Phillips-Schmidt-Shin
result = kpss(titanic_df['Fare'])
print(f"Kwiatkowski-Phillips-Schmidt-Shin Test: KPSS Statistic={result[0]}, p-value={result[1]}, Lags Used={result[2]}")
```

```
Augmented Dickey-Fuller Test: ADF Statistic=-20.554672335010277, p-value=0.0, Critical Values={'1%': -3.446129402876608, '5%':
-2.8684960761128346, '10%': -2.570475362616382}
Kwiatkowski-Phillips-Schmidt-Shin Test: KPSS Statistic=0.0682626545035462, p-value=0.1, Lags Used=6

C:\Users\squir\AppData\Local\Temp\ipykernel_7492\1209203104.py:7: InterpolationWarning: The test statistic is outside of the ra
nge of p-values available in the
look-up table. The actual p-value is greater than the p-value returned.

  result = kpss(titanic_df['Fare'])
```

In [23]:
```
# Parametric Statistical Hypothesis Tests
# Student's t-test
stat, p_value = ttest_ind(titanic_df['Age'], titanic_df['Fare'])
print(f"Student's t-test: t-statistic={stat}, p-value={p_value}")

# Paired Student's t-test
stat, p_value = ttest_rel(titanic_df['Age'], titanic_df['Fare'])
print(f"Paired Student's t-test: t-statistic={stat}, p-value={p_value}")

# Analysis of Variance Test (ANOVA)
result = f_oneway(titanic_df['Age'], titanic_df['Fare'])
print(f"Analysis of Variance Test (ANOVA): F-statistic={result.statistic}, p-value={result.pvalue}")
```

```
Student's t-test: t-statistic=-2.133594291454706, p-value=0.033167229221356787
Paired Student's t-test: t-statistic=-2.3116048159532463, p-value=0.021286109068868662
Analysis of Variance Test (ANOVA): F-statistic=4.552224600528116, p-value=0.03316722922137293
```

In [25]:
```
# Nonparametric Statistical Hypothesis Tests
# Mann-Whitney U Test
stat, p_value = mannwhitneyu(titanic_df['Age'], titanic_df['Fare'])
print(f"Mann-Whitney U Test: U-statistic={stat}, p-value={p_value}")

# Wilcoxon Signed-Rank Test
stat, p_value = wilcoxon(titanic_df['Age'], titanic_df['Fare'])
print(f"Wilcoxon Signed-Rank Test: W-statistic={stat}, p-value={p_value}")

# Kruskal-Wallis H Test
stat, p_value = kruskal(titanic_df['Age'], titanic_df['Fare'])
print(f"Kruskal-Wallis H Test: H-statistic={stat}, p-value={p_value}")

# Friedman Test
stat, p_value = friedmanchisquare(titanic_df['Age'], titanic_df['Fare'], titanic_df['Survived'])
print(f"Friedman Test: Chi2-statistic={stat}, p-value={p_value}")
```

```
Mann-Whitney U Test: U-statistic=116738.0, p-value=3.695066672640353e-17
Wilcoxon Signed-Rank Test: W-statistic=33145.0, p-value=2.275273810723764e-05
Kruskal-Wallis H Test: H-statistic=70.93572317835647, p-value=3.690546910949811e-17
Friedman Test: Chi2-statistic=646.7070828331326, p-value=3.709721088881058e-141
```

In [ ]: