Task 2

Name: Jaavanika L

1.Generate summary statistics (mean, median, std, etc.).

2.Create histograms and boxplots for numeric features.

3.Use pairplot/correlation matrix for feature relationships.

4.Identify patterns, trends, or anomalies in the data. 5.Make basic feature-level inferences from visuals

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

```
data = pd.read_csv("C:/Users/JAAVANIKA L/fall semester 22-23/Downloads/Titanic-Dataset.csv")
```

```
data.describe()
```

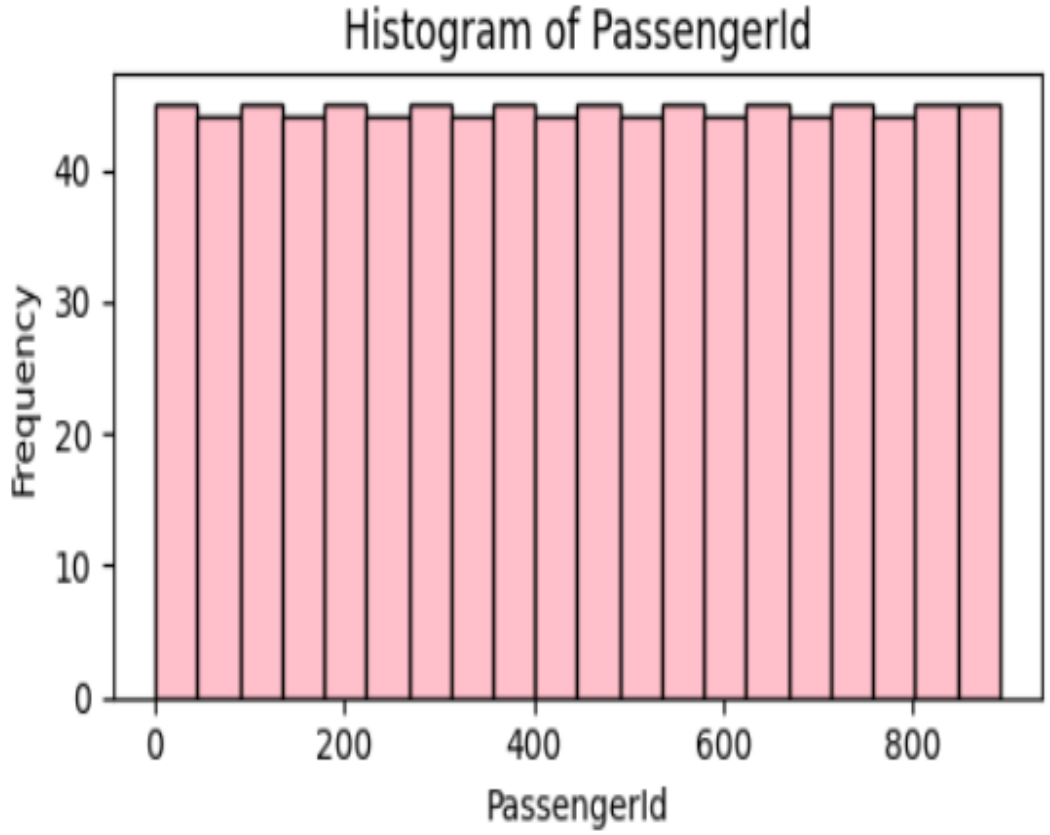
	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
count	891.000000	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	446.000000	0.383838	2.308642	29.699118	0.523008	0.381594	32.204208
std	257.353842	0.486592	0.836071	14.526497	1.102743	0.806057	49.693429
min	1.000000	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	223.500000	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	446.000000	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	668.500000	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

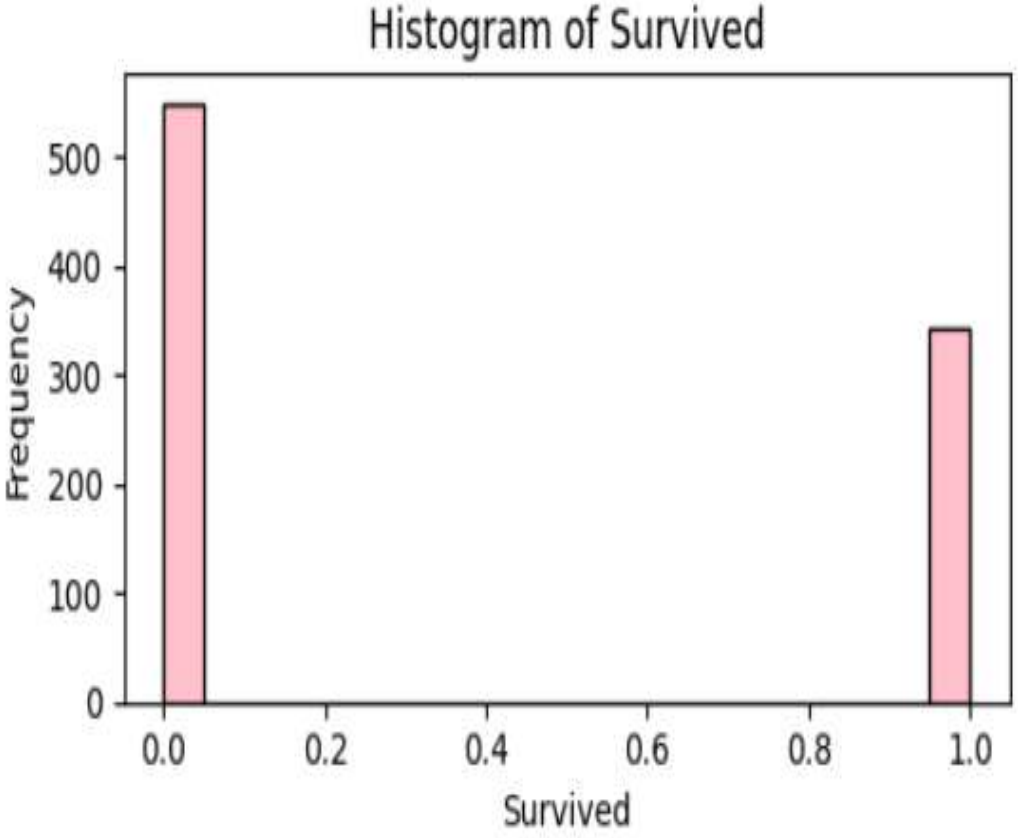
```
data.median(numeric_only=True)
```

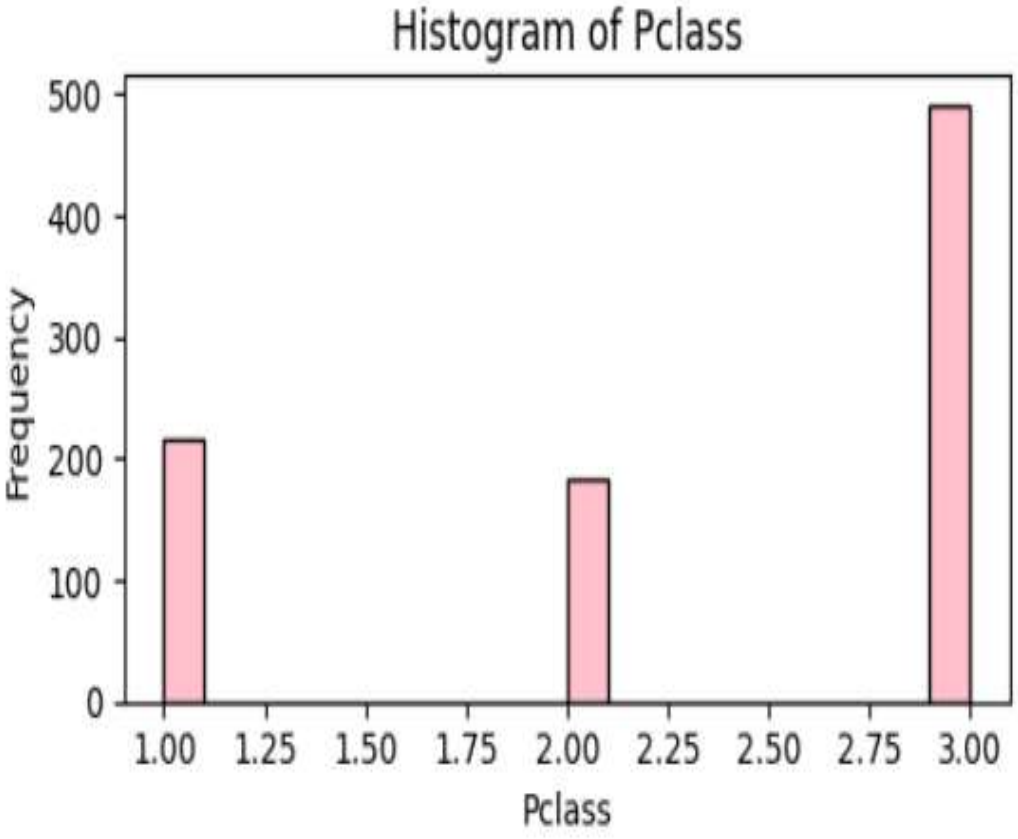
```
PassengerId    446.0000
Survived         0.0000
Pclass           3.0000
Age             28.0000
SibSp            0.0000
Parch            0.0000
Fare            14.4542
dtype: float64
```

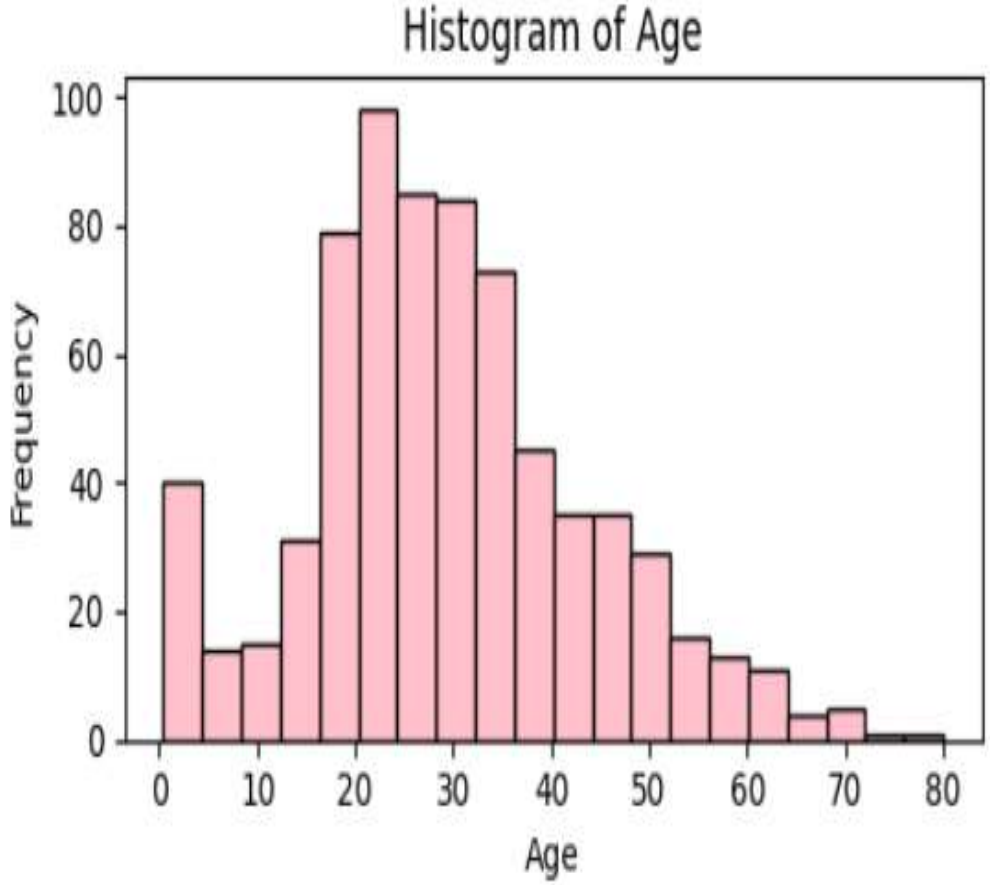
```
numeric_cols = data.select_dtypes(include=['int64', 'float64']).columns
```

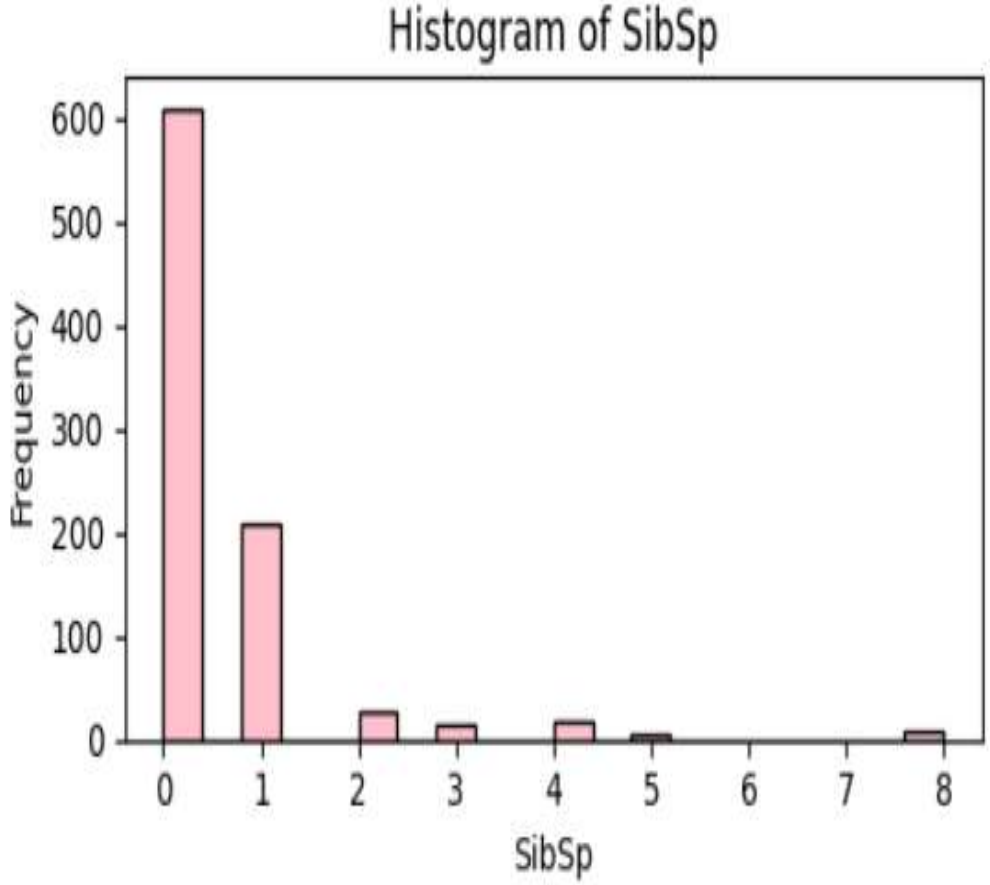
```
for col in numeric_cols:
    # Histogram
    plt.figure(figsize=(5, 3))
    plt.hist(data[col], bins=20, color='pink', edgecolor='black')
    plt.title(f'Histogram of {col}')
    plt.xlabel(col)
    plt.ylabel('Frequency')
    plt.tight_layout()
    plt.show()
```

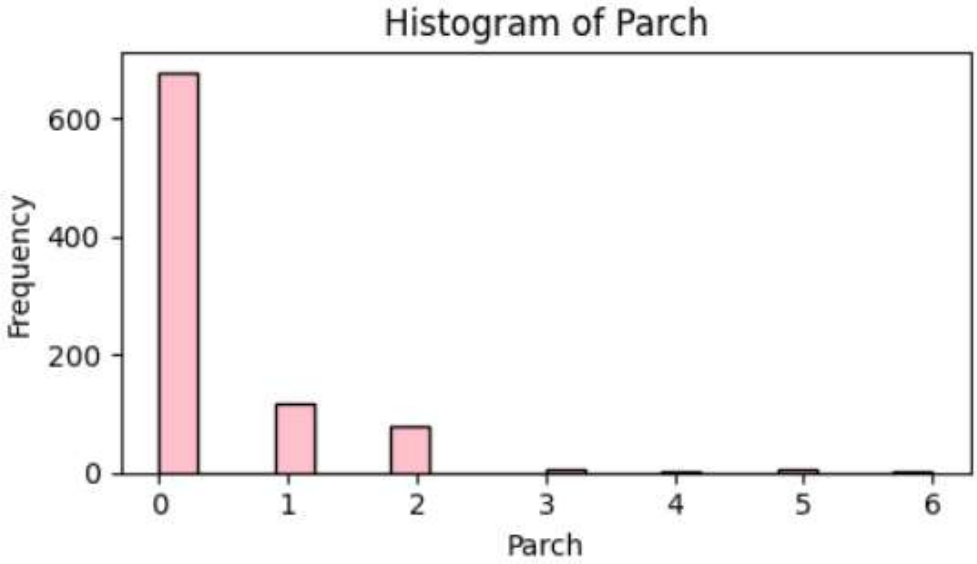


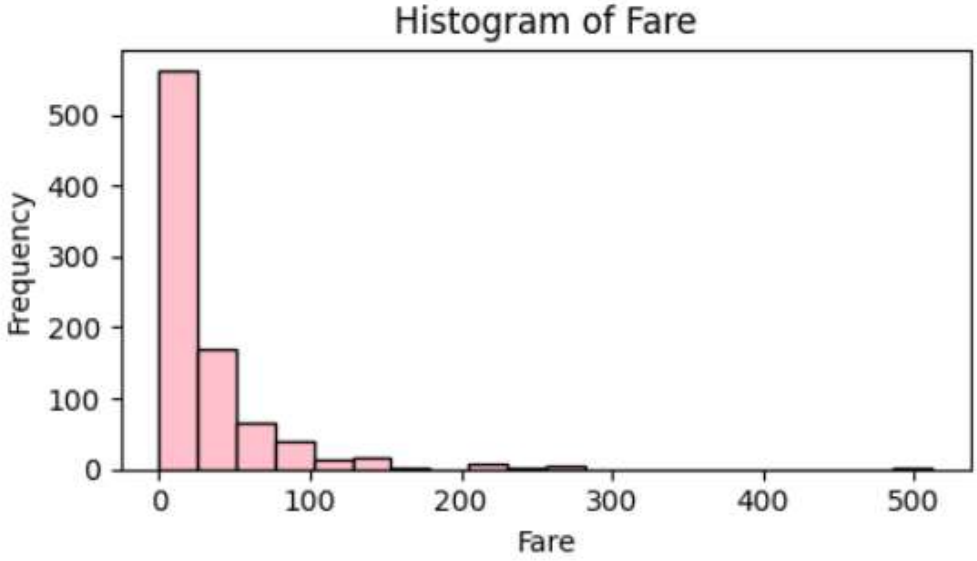

Histogram of Survived
Frequency
500
400
300
200
100
0
0.0
0.2
0.4
0.6
0.8
1.0
Survived


Histogram of Pclass
Frequency
500
400
300
200
100
0
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
Pclass

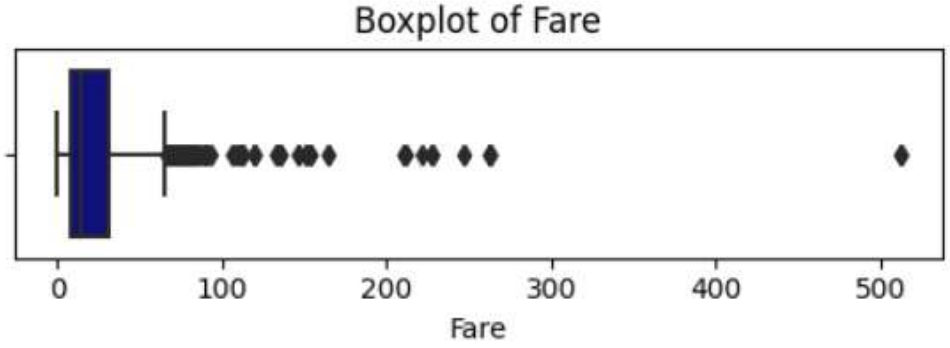

Histogram of Age
Frequency
Age
0
20
40
60
80
100
0
10
20
30
40
50
60
70
80


Histogram of SibSp
Frequency
SibSp
0
100
200
300
400
500
600
0
1
2
3
4
5
6
7
8

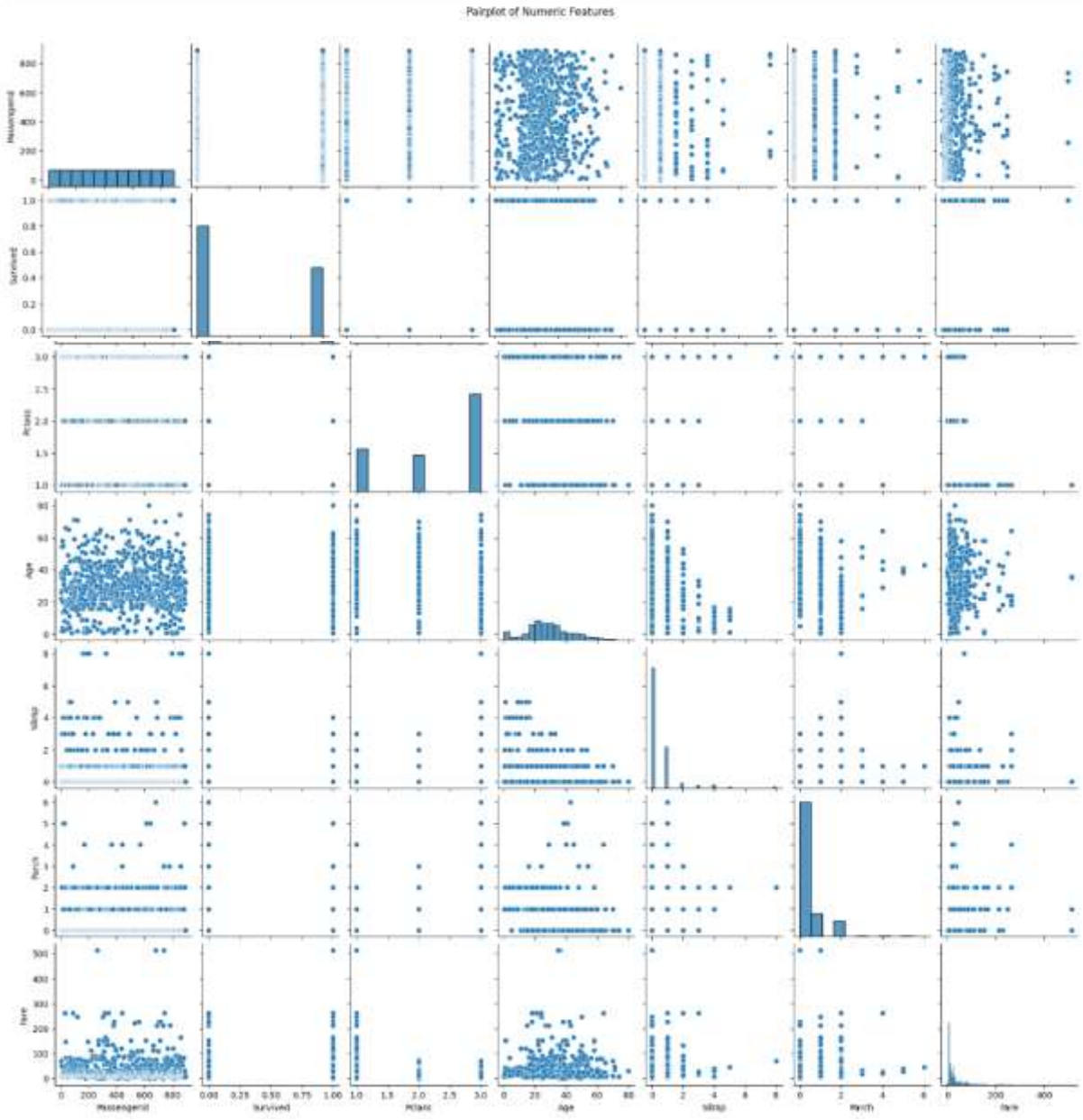





```
plt.figure(figsize=(5, 2))
sns.boxplot(x=data[col], color='darkblue')
plt.title(f'Boxplot of {col}')
plt.tight_layout()
plt.show()
```



```
sns.pairplot(data[numeric_cols])
plt.suptitle("Pairplot of Numeric Features", y=1.02)
plt.show()
```



```
plt.figure(figsize=(10, 6))
sns.heatmap(data[numeric_cols].corr(), annot=True, cmap='coolwarm', fmt=".2f")
plt.title("Correlation Matrix")
plt.show()
```