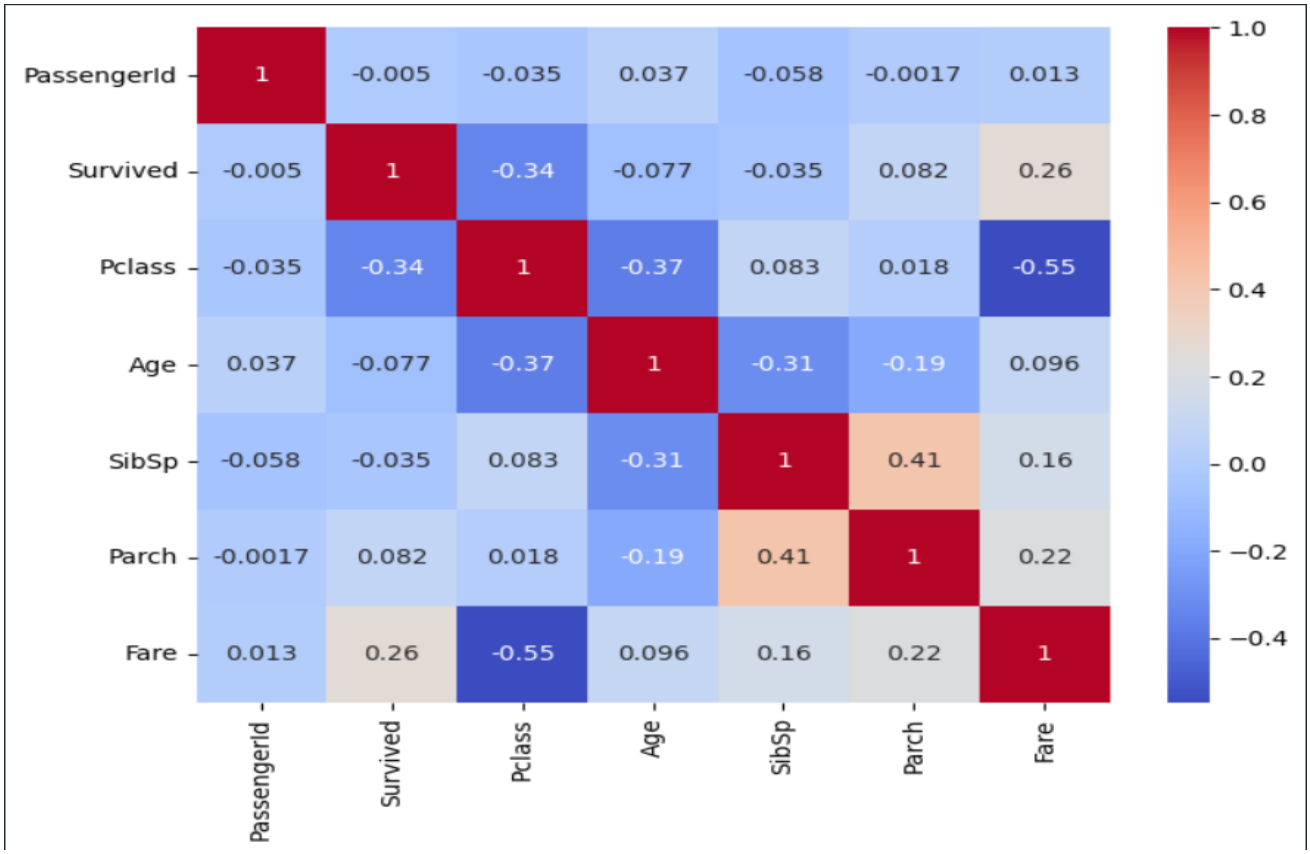
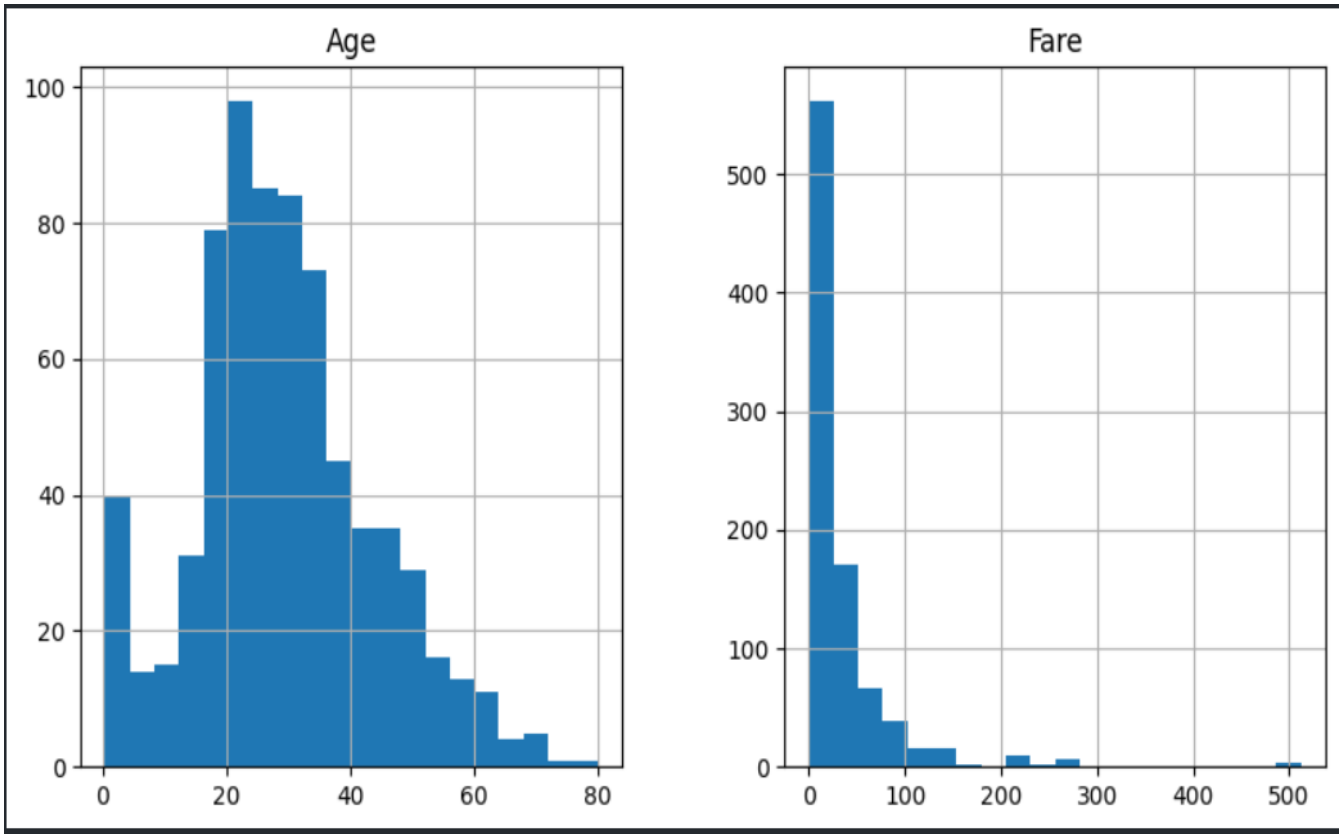
PDF report of findings:

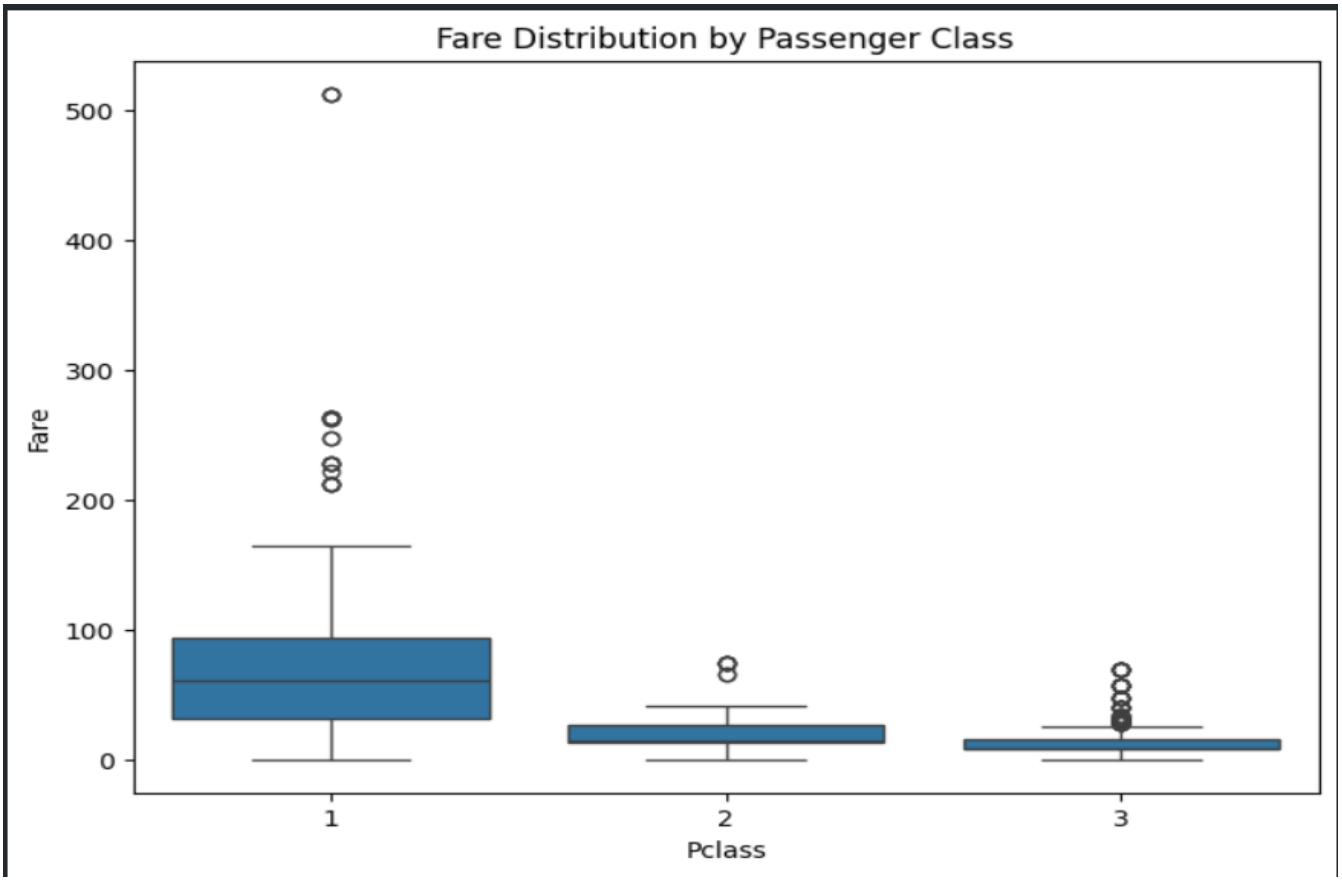
- **The training dataset contains records for 891 passengers.**
- **The overall survival rate was about 38%, meaning roughly 6 out of 10 passengers did not survive.**
- **Gender strongly influenced survival: around three-quarters of females survived, compared to fewer than one-fifth of males.**
- **Passenger class (Pclass) showed a clear pattern — survival chances were highest in 1st class and lowest in 3rd class.**
- **Fare values were positively related to survival — passengers who paid higher fares tended to survive more.**
- **Most passengers were between 20 and 40 years of age.**
- **Children under 10 years old showed higher survival rates compared to adults.**
- **Older passengers (above 60 years) were far less likely to survive.**
- **Family size had mixed effects: survival was better for passengers with 1–2 relatives onboard, while those traveling alone or in large groups had lower chances.**
- **The correlation heatmap confirmed strong links between survival and features such as passenger class, fare, and gender.**
- **The majority of passengers boarded at Southampton (S), followed by Cherbourg (C) and Queenstown (Q).**
- **Passengers from Cherbourg (C) generally had higher survival rates than those from Southampton or Queenstown.**
- **Boxplots of fare across classes highlighted that 1st class passengers paid significantly higher fares and had greater survival chances.**
- **The scatterplot of Age vs Fare revealed that young, high-fare passengers (especially children in 1st class) were more likely to survive.**
- **In summary, gender, class, and fare were the strongest predictors of survival, age played a moderate role, and the port of embarkation added some influence.**

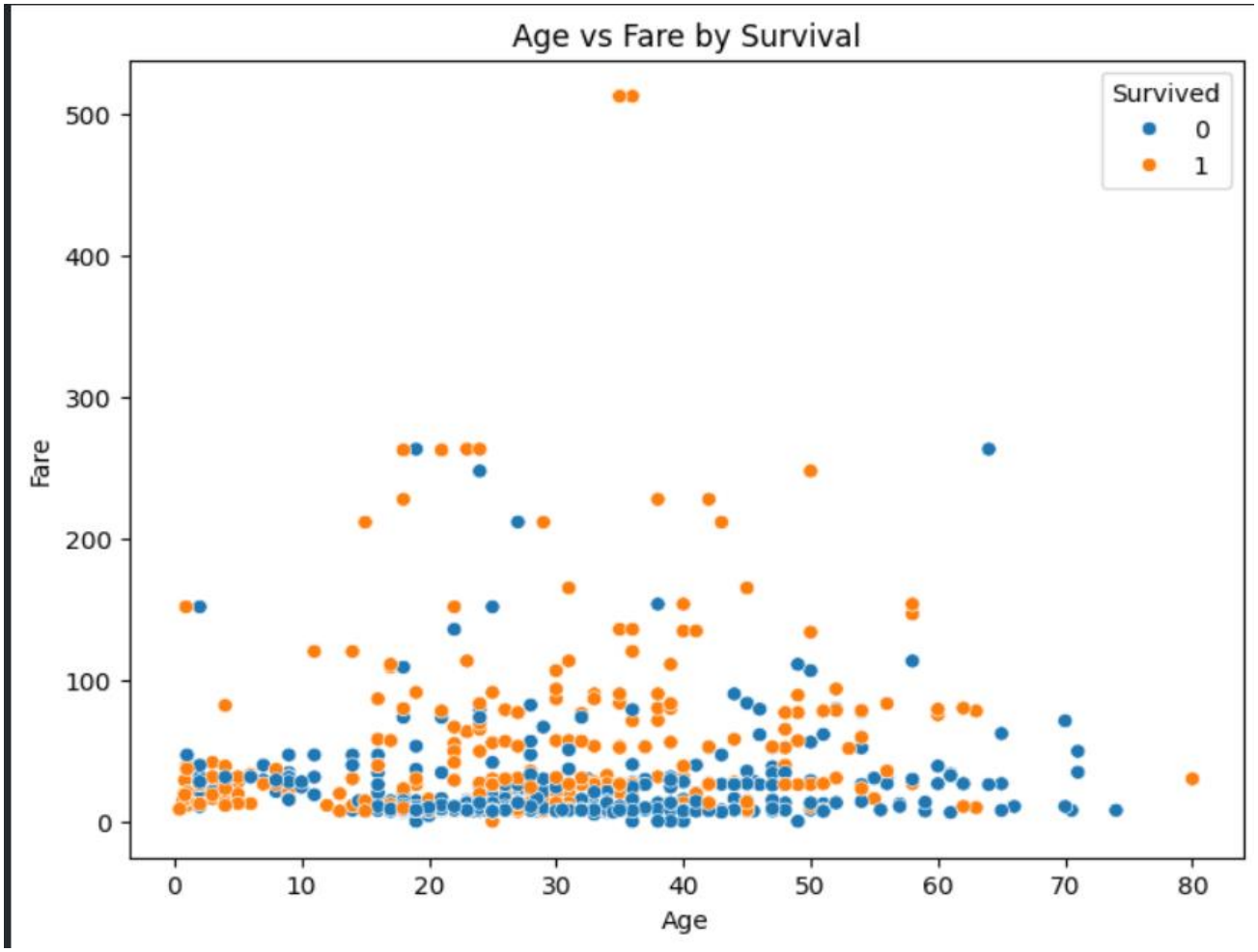
```
       PassengerId    Survived      Pclass         Age       SibSp  \
count   891.000000  891.000000  891.000000  714.000000  891.000000
mean    446.000000    0.383838    2.308642   29.699118    0.523008
std     257.353842    0.486592    0.836071   14.526497    1.102743
min       1.000000    0.000000    1.000000    0.420000    0.000000
25%     223.500000    0.000000    2.000000   20.125000    0.000000
50%     446.000000    0.000000    3.000000   28.000000    0.000000
75%     668.500000    1.000000    3.000000   38.000000    1.000000
max     891.000000    1.000000    3.000000   80.000000    8.000000

             Parch        Fare
count  891.000000  891.000000
mean     0.381594   32.204208
std      0.806057   49.693429
min      0.000000    0.000000
25%      0.000000    7.910400
50%      0.000000   14.454200
75%      0.000000   31.000000
max      6.000000  512.329200
Sex
male      577
female    314
Name: count, dtype: int64
Survived
0    549
1    342
Name: count, dtype: int64
Pclass
3    491
1    216
2    184
Name: count, dtype: int64
```

```
   PassengerId  Survived  Pclass  \
0            1         0       3
1            2         1       1
2            3         1       3
3            4         1       1
4            5         0       3

                                                Name     Sex   Age  SibSp  \
0                            Braund, Mr. Owen Harris    male  22.0      1
1  Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1
2                             Heikkinen, Miss. Laina  female  26.0      0
3       Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1
4                           Allen, Mr. William Henry    male  35.0      0

   Parch            Ticket     Fare Cabin Embarked
0      0         A/5 21171   7.2500   NaN        S
1      0          PC 17599  71.2833   C85        C
2      0  STON/O2. 3101282   7.9250   NaN        S
3      0            113803  53.1000  C123        S
4      0            373450   8.0500   NaN        S
Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',
       'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],
      dtype='object')
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   PassengerId  891 non-null    int64
 1   Survived     891 non-null    int64
 2   Pclass       891 non-null    int64
 3   Name         891 non-null    object
 4   Sex          891 non-null    object
 5   Age          714 non-null    float64
 6   SibSp        891 non-null    int64
 7   Parch        891 non-null    int64
 8   Ticket       891 non-null    object
 9   Fare         891 non-null    float64
 10  Cabin        204 non-null    object
 11  Embarked     889 non-null    object
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
None
```


PassengerId
Survived
Pclass
Age
SibSp
Parch
Fare
1 -0.005 -0.035 0.037 -0.058 -0.0017 0.013
-0.005 1 -0.34 -0.077 -0.035 0.082 0.26
-0.035 -0.34 1 -0.37 0.083 0.018 -0.55
0.037 -0.077 -0.37 1 -0.31 -0.19 0.096
-0.058 -0.035 0.083 -0.31 1 0.41 0.16
-0.0017 0.082 0.018 -0.19 0.41 1 0.22
0.013 0.26 -0.55 0.096 0.16 0.22 1
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4


Age
100
80
60
40
20
0
0
20
40
60
80
Fare
500
400
300
200
100
0
0
100
200
300
400
500


Fare Distribution by Passenger Class
Fare
Pclass
0
100
200
300
400
500
1
2
3


Age vs Fare by Survival
Survived
0
1
Fare
Age
0
100
200
300
400
500
0
10
20
30
40
50
60
70
80