

TITANIC SURVIVAL ANALYSIS

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

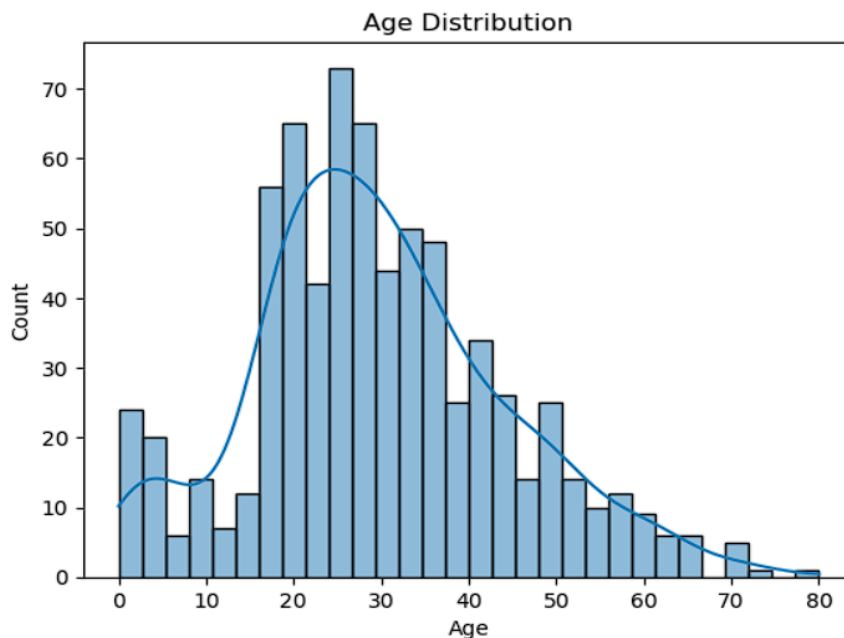
```
%matplotlib inline
```

```
df = pd.read_csv("titanic_train.csv")
df.head()
```

	Passenger_Id	Survived	Ticket_class	Gender	Age	Siblings_Spouse_Aboard	Parents_Children_Aboard	Fare	Embarked_port
0	1	No	3	male	22.0	1	0	7.25	Southampton
1	2	Yes	1	female	38.0	1	0	71.28	Cherbourg
2	3	Yes	3	female	26.0	0	0	7.93	Southampton
3	4	Yes	1	female	35.0	1	0	53.10	Southampton
4	5	No	3	male	35.0	0	0	8.05	Southampton

1. Age Distribution

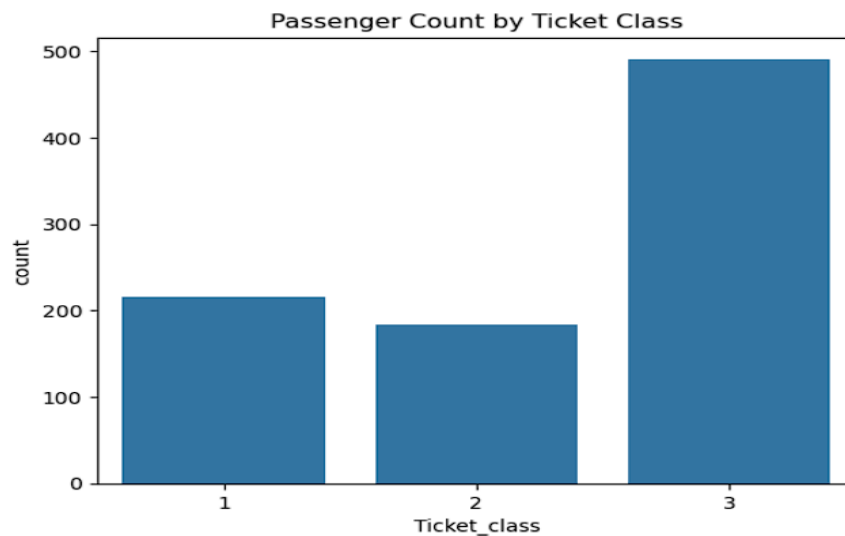
```
[4]: sns.histplot(df['Age'], bins=30, kde=True)
plt.title('Age Distribution')
plt.show()
```



Most passengers were between 20–40 years old. A few infants and elderly passengers were also on board, but the majority were young adults.

2. Distribution of passengers across different ticket classes

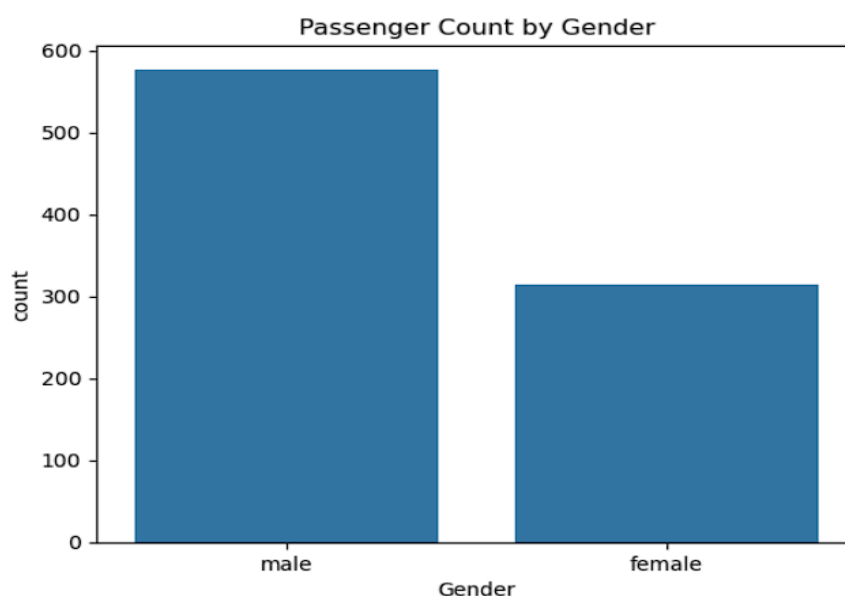
```
[5]: sns.countplot(x='Ticket_class', data=df)
plt.title('Passenger Count by Ticket Class')
plt.show()
```



Most passengers travelled in the 3rd class, followed by 1st and 2nd. This indicates that the Titanic served a large number of economically lower-class travellers.

3. Passenger count by Gender

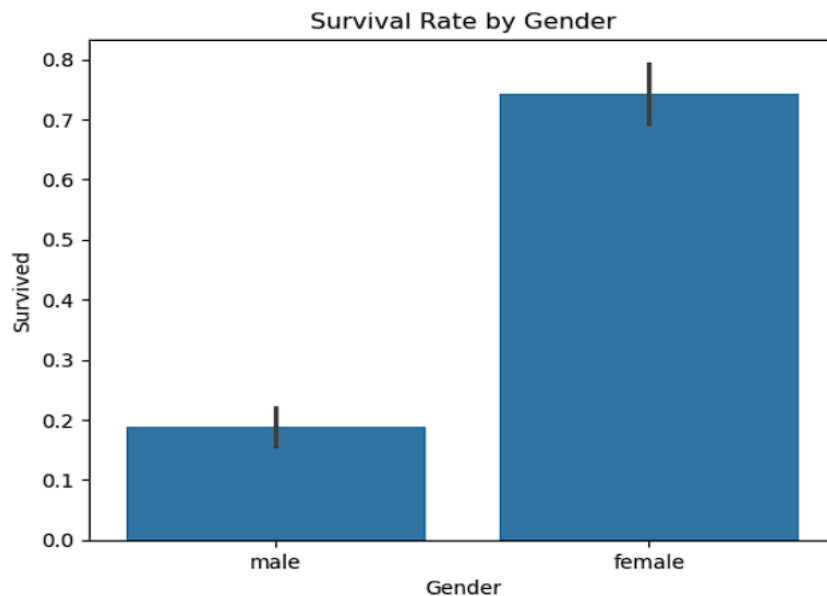
```
[6]: sns.countplot(x='Gender', data=df)
plt.title('Passenger Count by Gender')
plt.show()
```



There were significantly more male passengers than female. This imbalance is important context when analysing gender-based survival patterns.

4. Plot survival rate by gender

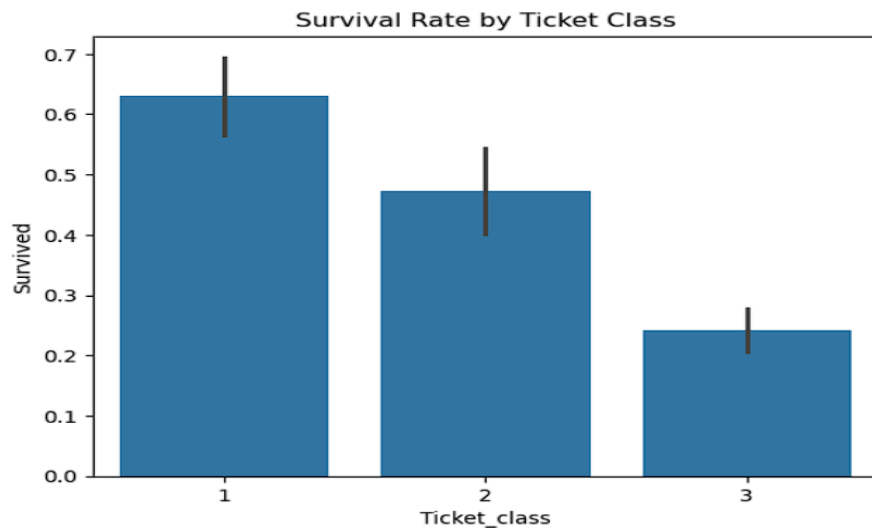
```
[7]: sns.barplot(x='Gender', y='Survived', data=df)
plt.title('Survival Rate by Gender')
plt.show()
```



Females had a significantly higher survival count compared to males. This supports the historical fact that women were prioritized during evacuation.

5. Survival rate by ticket class

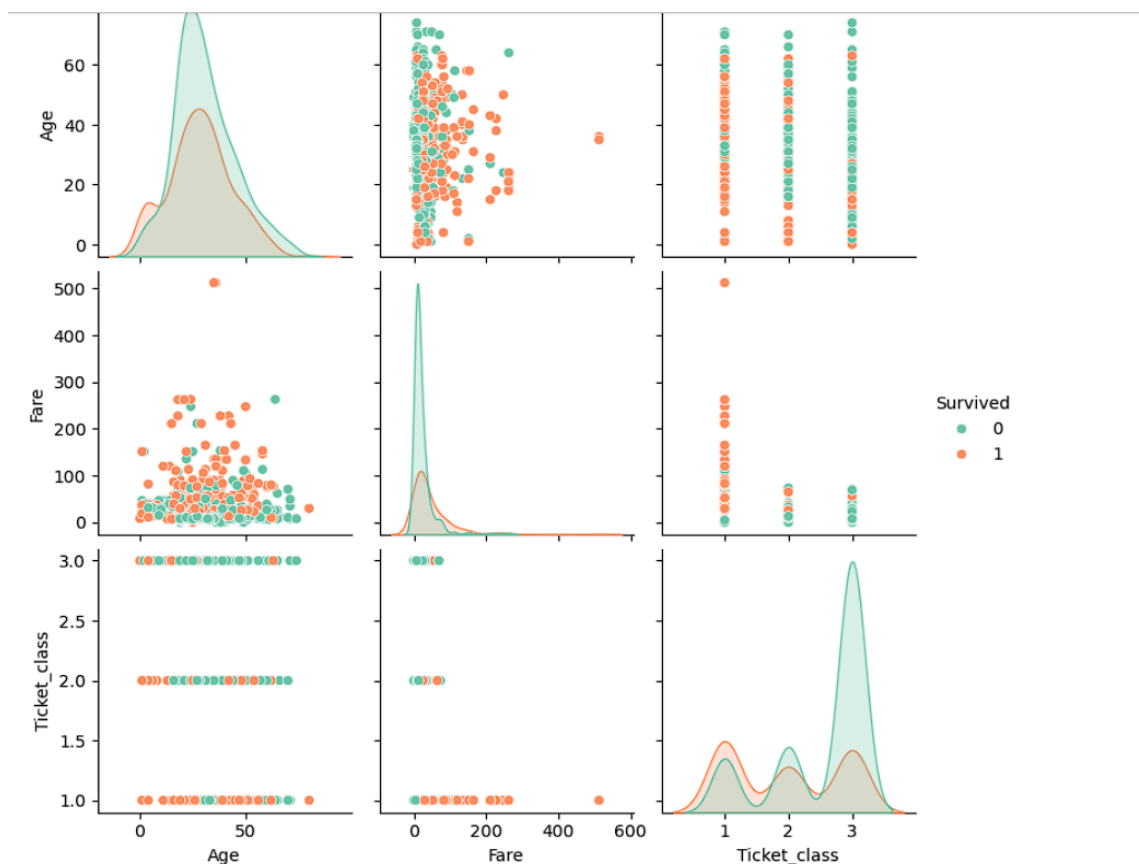
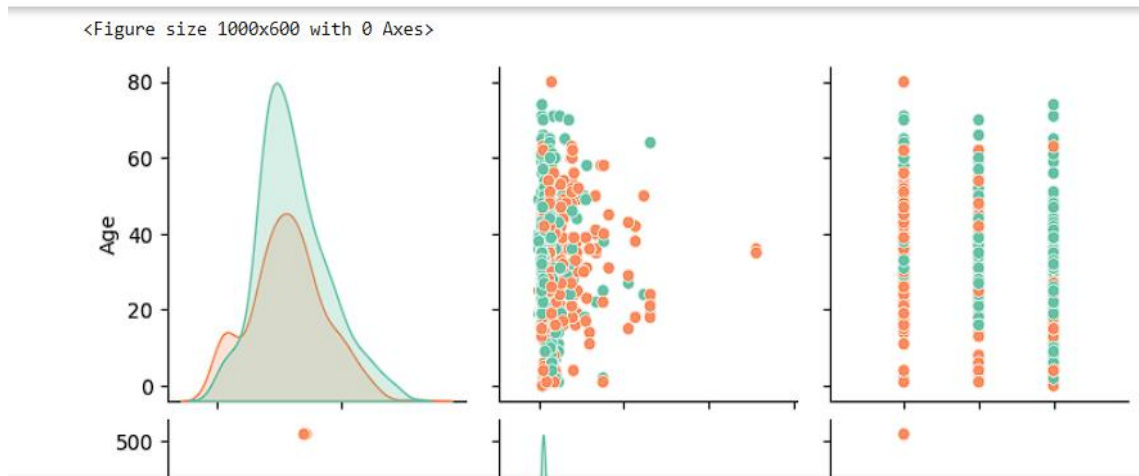
```
[8]: sns.barplot(x='Ticket_class', y='Survived', data=df)
plt.title('Survival Rate by Ticket Class')
plt.show()
```



Passengers in 1st class had the highest survival rates, while 3rd class passengers had the lowest. This reflects social inequality in access to lifeboats or safety.

6. Pairplot to see Relationships

```
[16]: sns.pairplot(df[selected_cols], hue='Survived', diag_kind='kde', palette='Set2')  
plt.show()
```



Survivors are generally younger, paid higher fares, and were more likely to be in 1st class. The Pairplot visually confirms clusters that relate to higher survival chances.