

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

sns.set(style='whitegrid')
%matplotlib inline
```



```
df = pd.read_csv('Downloads/train.csv')  
df.head()
```



Show hidden output



```
# Dataset Overview  
print('Shape:', df.shape)  
df.info()  
df.describe()  
df.isnull().sum()
```



```
# Univariate Analysis  
sns.countplot(x='Survived', data=df)  
plt.title('Survival Count')  
plt.show()
```



```
# Bivariate Analysis  
sns.boxplot(x='Survived', y='Age', data=df)  
plt.title('Age Distribution by Survival')  
plt.show()
```

```
# Correlation Heatmap
plt.figure(figsize=(10, 8))
sns.heatmap(df.corr(), annot=True, cmap='coolwarm')
plt.title('Feature Correlation')
plt.show()
```

[2]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

```

Shape: (891, 12)
<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

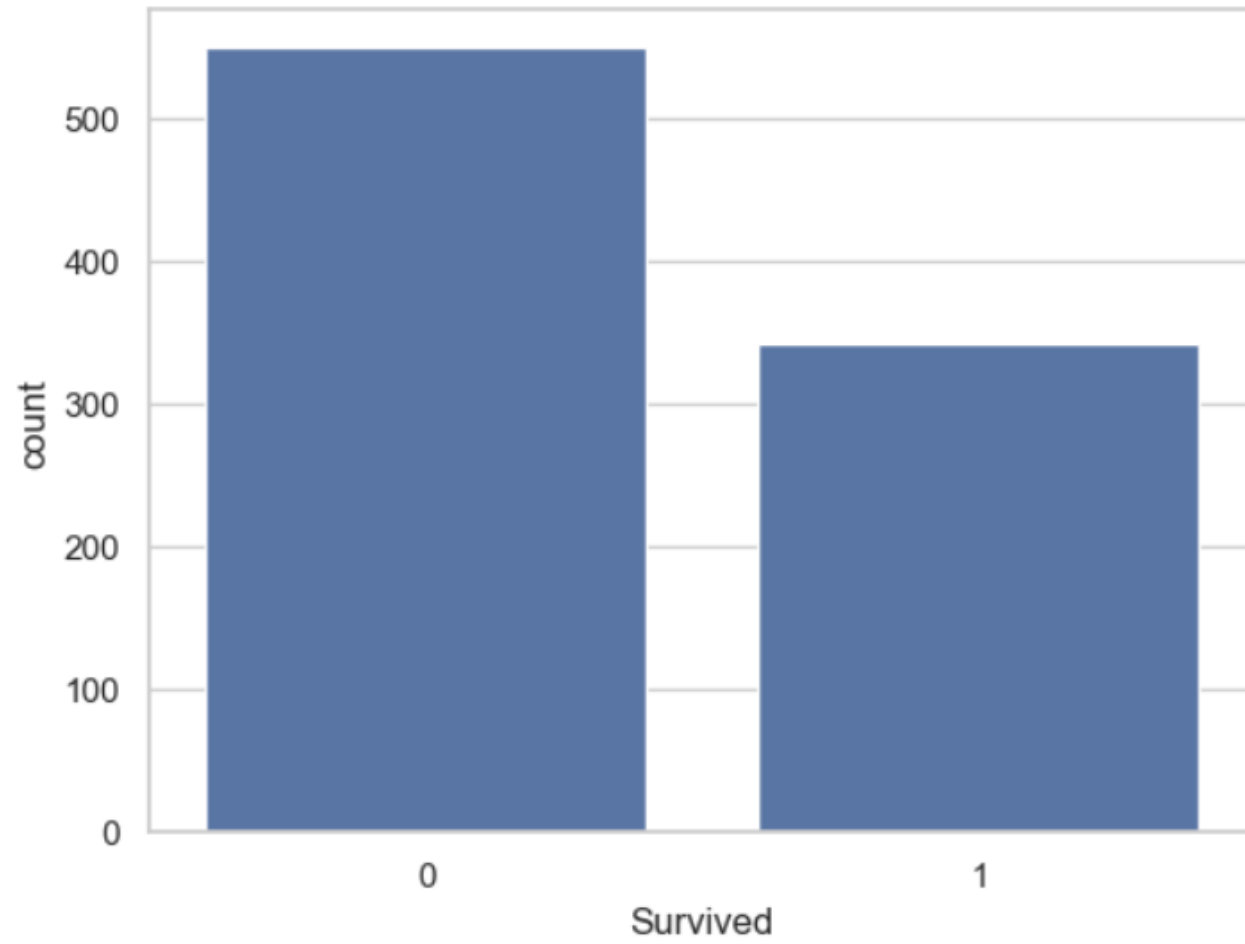
```

```

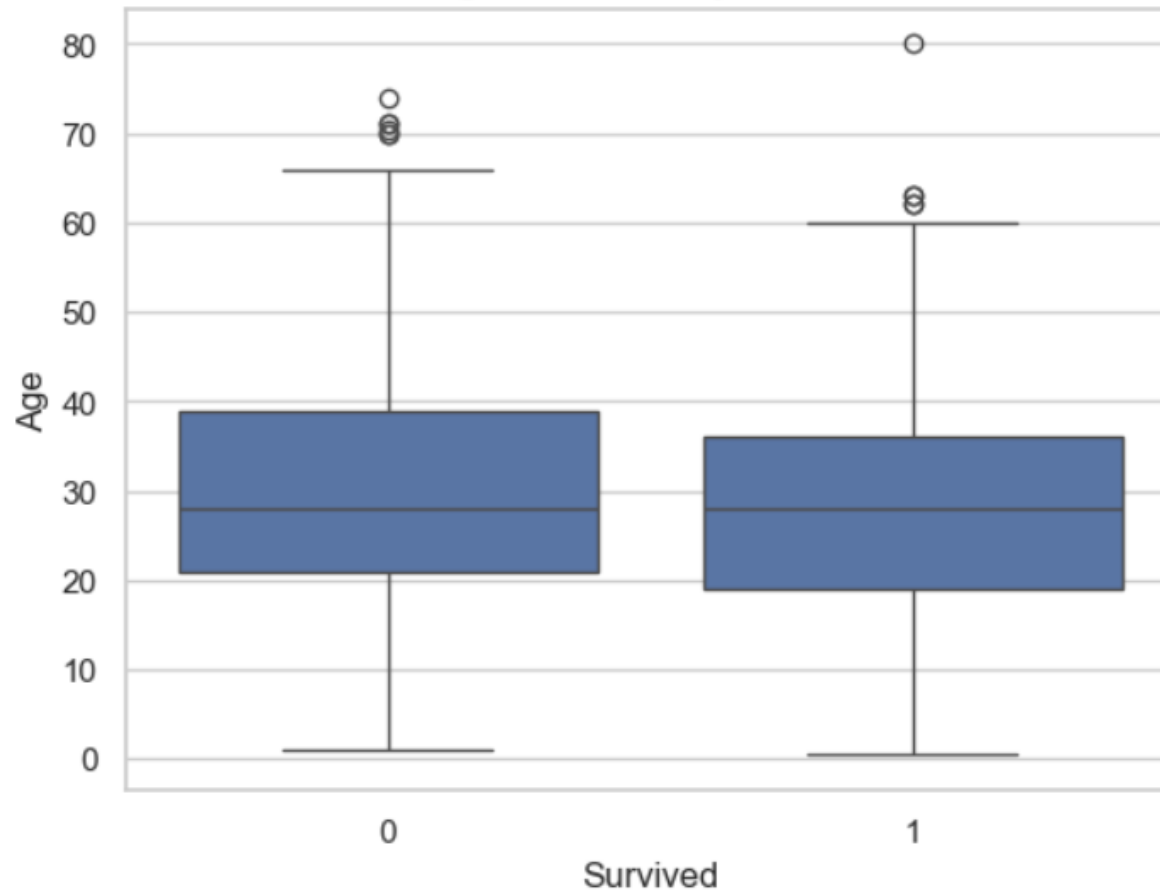
3]: PassengerId      0
   Survived         0
   Pclass           0
   Name             0
   Sex              0
   Age             177
   SibSp           0
   Parch           0
   Ticket          0
   Fare            0
   Cabin          687
   Embarked        2
dtypes: int64(5), object(7)

```

Survival Count



Age Distribution by Survival





Summary of Findings

- Majority did not survive
- Age and Fare show variability across survival
- Many missing values in Age and Cabin
- Pclass, Sex, and Fare appear to have notable influence on survival