

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

df = pd.read_csv('titanic dataset.csv')
df.head()
```

	PassengerId	Survived
0	892	0
1	893	1
2	894	0
3	895	0
4	896	1

SUMMARIZING THE DATA

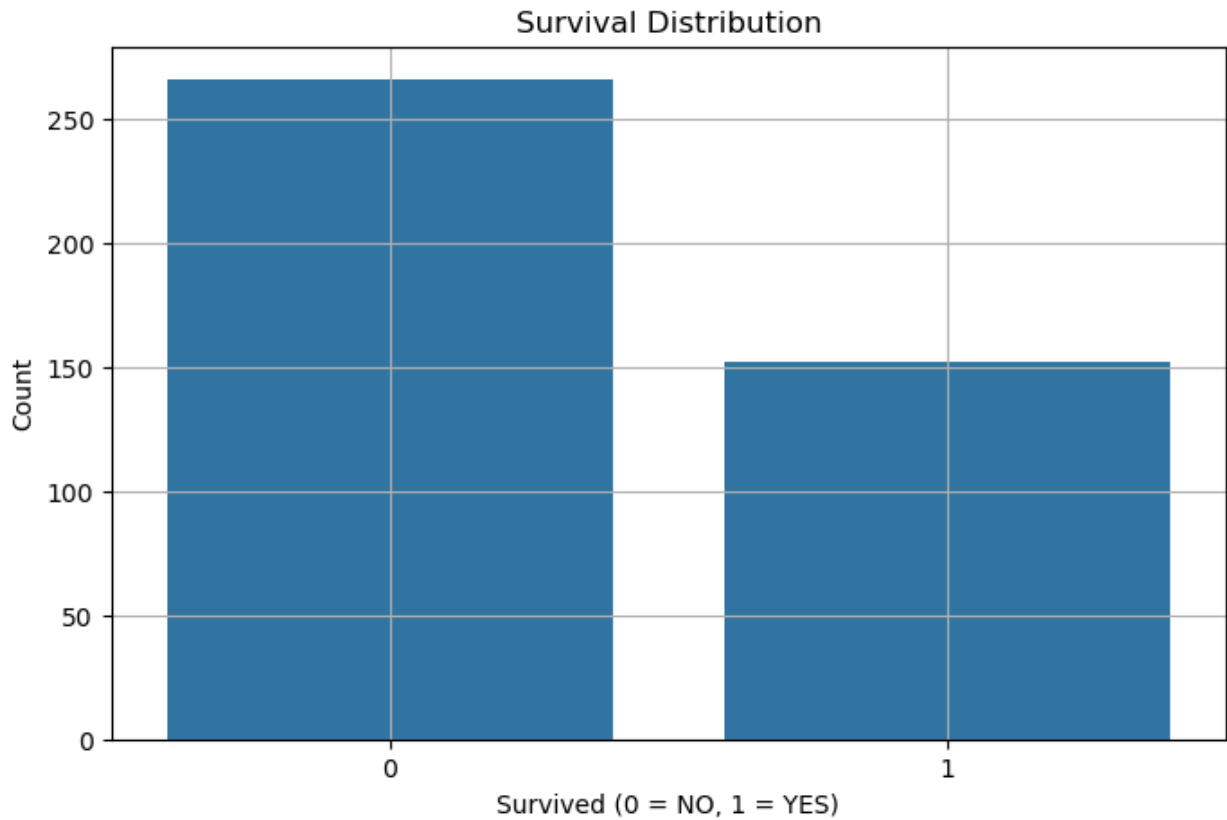
```
df.describe(include = 'all')
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 418 entries, 0 to 417
Data columns (total 2 columns):
#   Column          Non-Null Count  Dtype
---  -
0   PassengerId     418 non-null   int64
1   Survived        418 non-null   int64
dtypes: int64(2)
memory usage: 6.7 KB
```

EXPLORATORY DATA ANALYSIS

1. SURVIVAL DISTRIBUTION

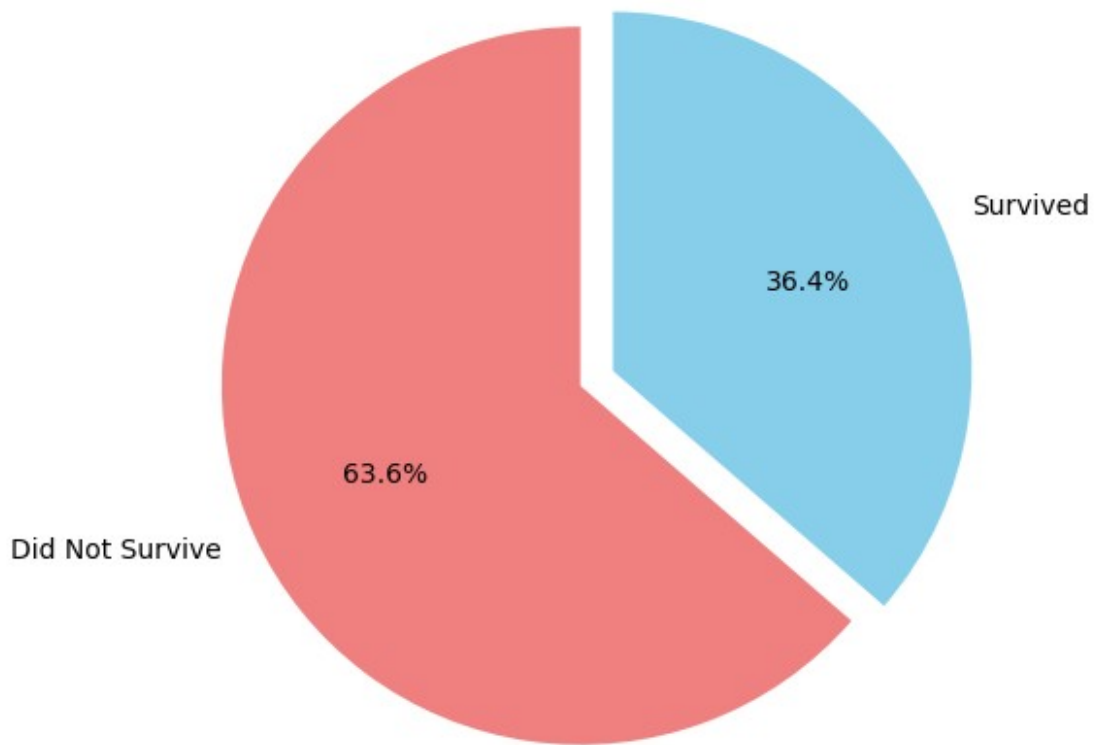
```
plt.figure(figsize=(8,5))
sns.countplot(x = 'Survived', data = df)
plt.title('Survival Distribution')
plt.xlabel('Survived (0 = NO, 1 = YES)')
plt.ylabel('Count')
plt.grid(True)
plt.show()
```



SURVIVAL PROPORTIONS

```
plt.figure(figsize=(8, 6))
df['Survived'].value_counts().plot.pie(
    labels=['Did Not Survive', 'Survived'],
    autopct='%1.1f%%',
    startangle=90,
    colors=['lightcoral', 'skyblue'],
    explode=[0, 0.1]
)
plt.title('Survival Proportions', fontsize=16)
plt.ylabel('')
plt.show()
```

Survival Proportions



By analyzing the above steps, the titanic survival rate is less i.e., 36.4% as compared to did not survived i.e., 63.6%. The count of Survived passengers is 255 and 151 for not survived.