

```
# Import libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Load the Titanic dataset
url = "/content/drive/MyDrive/ProdigyInfotech/train.csv"
titanic = pd.read_csv(url)

# Display the first few rows of the dataset
print(titanic.head())

# Check for missing values
print(titanic.isnull().sum())

# Impute missing values
titanic['Age'].fillna(titanic['Age'].median(), inplace=True)
titanic['Embarked'].fillna(titanic['Embarked'].mode()[0], inplace=True)
titanic.drop('Cabin', axis=1, inplace=True) # Drop Cabin column

# Explore the relationships between variables
# Let's start with a pair plot to visualize relationships between numerical variables
sns.pairplot(titanic, hue='Survived', diag_kind='kde')
plt.show()

# Explore the distribution of categorical variables
plt.figure(figsize=(12, 6))

# Plot the distribution of passenger class
plt.subplot(2, 3, 1)
sns.countplot(x='Pclass', hue='Survived', data=titanic)

# Plot the distribution of sex
plt.subplot(2, 3, 2)
sns.countplot(x='Sex', hue='Survived', data=titanic)

# Plot the distribution of embarked port
plt.subplot(2, 3, 3)
sns.countplot(x='Embarked', hue='Survived', data=titanic)

# Plot the distribution of siblings/spouses aboard
plt.subplot(2, 3, 4)
```

```
plt.subplot(2, 3, 4)
sns.countplot(x='SibSp', hue='Survived', data=titanic)

# Plot the distribution of parents/children aboard
plt.subplot(2, 3, 5)
sns.countplot(x='Parch', hue='Survived', data=titanic)

plt.tight_layout()
plt.show()
```



	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	Cummings, 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

```

PassengerId      0
Survived          0
Pclass           0
Name             0
Sex              0
Age             177
SibSp            0
Parch            0
Ticket           0
Fare             0
Cabin           687
Embarked         2
dtype: int64

```





