In [ ]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import matplotlib

matplotlib.use('TkAgg')

df = pd.read_csv('train.csv')

print(df.head())
print(df.info())
print(df.describe())
print(df.columns)

print(df.isnull().sum())


sns.heatmap(df.isnull(), cbar=False, cmap='viridis')
plt.title("Missing Value Heatmap")
plt.show()
plt.close()

df['Age'].hist(bins=30, edgecolor='black')
plt.title('Age Distribution')
plt.xlabel('Age')
plt.ylabel('Count')
plt.show()
plt.close()

df['Fare'].hist(bins=40)
plt.title('Fare Distribution')
plt.xlabel('Fare')
plt.ylabel('Count')
plt.show()
plt.close()

df['Survived'].value_counts().plot(kind='bar')
plt.title('Survival Count')
plt.xlabel('Survived (0 = No, 1 = Yes)')
plt.ylabel('Count')
plt.show()
plt.close()


sns.countplot(x='Pclass', data=df)
plt.title("Passenger Class Count")
plt.show()
plt.close()

sns.countplot(x='Survived', hue='Sex', data=df)
plt.title("Survival by Gender")
plt.show()
plt.close()

sns.countplot(x='Survived', hue='Pclass', data=df)
plt.title("Survival by Class")
plt.show()
```

```
plt.close()

sns.boxplot(x='Survived', y='Fare', data=df)
plt.title("Fare Distribution by Survival")
plt.show()
plt.close()

numeric_df = df.select_dtypes(include=[float, int])

sns.heatmap(numeric_df.corr(), annot=True, cmap='coolwarm')
plt.title("Correlation Matrix")
plt.show()
plt.close()

sns.pairplot(df[['Age', 'Fare', 'Pclass', 'Survived']])
plt.show()
plt.close()
```