In [ ]:

```python
# Importing Libraries
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Load the Titanic dataset
titanic = sns.load_dataset('titanic')

# 1. Basic Information
print("=== Dataset Info ===")
print(titanic.info())

print("\n=== First 5 Rows ===")
print(titanic.head())

# 2. Statistical Summary
print("\n=== Statistical Summary ===")
print(titanic.describe(include='all'))

# 3. Checking Missing Values
print("\n=== Missing Values ===")
print(titanic.isnull().sum())

# 4. Value Counts for Important Categorical Variables
print("\n=== Value Counts ===")
print("\nSex:\n", titanic['sex'].value_counts())
print("\nClass:\n", titanic['class'].value_counts())
print("\nEmbark Town:\n", titanic['embark_town'].value_counts())

# ============================
# Univariate Analysis
# ============================

# 5. Age Distribution
plt.figure(figsize=(8,5))
sns.histplot(titanic['age'].dropna(), kde=True)
plt.title('Age Distribution')
plt.xlabel('Age')
plt.ylabel('Frequency')
```

```
plt.show()
print("Observation: Most passengers are between 20 to 40 years old.")

# 6. Sex Count
plt.figure(figsize=(6,4))
sns.countplot(data=titanic, x='sex')
plt.title('Gender Distribution')
plt.show()
print("Observation: There are more males than females on board.")

# 7. Passenger Class Count
plt.figure(figsize=(6,4))
sns.countplot(data=titanic, x='class')
plt.title('Passenger Class Distribution')
plt.show()
print("Observation: Most passengers are from Third Class.")

# ============================
# Bivariate Analysis
# ============================

# 8. Survival Rate by Sex
plt.figure(figsize=(6,4))
sns.barplot(data=titanic, x='sex', y='survived')
plt.title('Survival Rate by Gender')
plt.show()
print("Observation: Females had a higher survival rate compared to males.")

# 9. Survival Rate by Class
plt.figure(figsize=(6,4))
sns.barplot(data=titanic, x='class', y='survived')
plt.title('Survival Rate by Class')
plt.show()
print("Observation: First Class passengers had higher chances of survival.")

# 10. Survival Rate by Embark Town
plt.figure(figsize=(8,5))
sns.barplot(data=titanic, x='embark_town', y='survived')
plt.title('Survival Rate by Embark Town')
plt.show()
print("Observation: Passengers from Cherbourg had the highest survival rate.")
```

```python
# ============================
# Multivariate Analysis
# ============================

# 11. Pairplot
sns.pairplot(titanic.dropna(), hue='survived', vars=['age', 'fare', 'parch', 'sibsp'])
plt.suptitle('Pairplot of Age, Fare, Parch, SibSp', y=1.02)
plt.show()
print("Observation: Survivors tend to have slightly higher fare and fewer siblings/parents.")

# 12. Correlation Heatmap
plt.figure(figsize=(10,8))
sns.heatmap(titanic.corr(), annot=True, cmap='coolwarm', linewidths=0.5)
plt.title('Correlation Heatmap')
plt.show()
print("Observation: Fare and Pclass are negatively correlated. Survival is positively correlated with Fare.")

# ============================
# Other Visualizations
# ============================

# 13. Boxplot - Age vs Class
plt.figure(figsize=(8,5))
sns.boxplot(data=titanic, x='class', y='age')
plt.title('Age Distribution across Classes')
plt.show()
print("Observation: First-class passengers are generally older than third-class passengers.")

# 14. Scatterplot - Fare vs Age
plt.figure(figsize=(8,5))
sns.scatterplot(data=titanic, x='age', y='fare', hue='survived')
plt.title('Fare vs Age Scatterplot')
plt.show()
print("Observation: Higher fares are often associated with survivors, regardless of age.")

# 15. Histogram - Fare Distribution
plt.figure(figsize=(8,5))
sns.histplot(titanic['fare'], bins=30, kde=True)
plt.title('Fare Distribution')
plt.show()
```

```
print("Observation: Most passengers paid lower fares, but some outliers paid extremely high fares.")

# ============================
# Final Summary
# ============================

print("""
=== Summary of Findings ===
1. Majority of passengers were young adults (20-40 years old).
2. There were more males than females, but females had a higher surviva l rate.
3. First-class passengers had a better survival rate compared to low.
""")
```