Here's a detailed explanation of the code and potential viva questions with answers:

---

### \*\*Code Explanation\*\*

```python

import seaborn as sns

import matplotlib.pyplot as plt

# Load Titanic dataset and clean data

titanic = sns.load\_dataset('titanic')

titanic = titanic.dropna(subset=['age', 'sex', 'survived'])

# Boxplot

plt.figure(figsize=(8, 6))

sns.boxplot(x='sex', y='age', hue='survived', data=titanic)

plt.title('Box Plot of Age by Sex and Survival Status')

plt.xlabel('Sex')

plt.ylabel('Age')

plt.legend(title='Survived', labels=['No', 'Yes'])

plt.show()

# Violinplot

plt.figure(figsize=(8, 6))

sns.violinplot(x='sex', y='age', hue='survived', data=titanic, split=True)

plt.title('Violin Plot of Age by Sex and Survival Status')

plt.xlabel('Sex')

plt.ylabel('Age')

plt.legend(title='Survived', labels=['No', 'Yes'])

plt.show()

```

---

### \*\*Viva Questions & Answers\*\*

#### \*\*Basic Concepts\*\*

\*\*Q1: What does `dropna()` do here?\*\*

A1: Removes rows with missing values in the `age`, `sex`, or `survived` columns to ensure clean data for plotting.

\*\*Q2: What does `hue='survived'` do?\*\*

A2: Splits the data by survival status (0/1) to compare age distributions between survivors and non-survivors.

\*\*Q3: Why use both boxplots and violinplots?\*\*

A3:

- \*\*Boxplot\*\*: Shows quartiles, median, and outliers

- \*\*Violinplot\*\*: Adds density estimation to show distribution shape

---

#### \*\*Plot Interpretation\*\*

\*\*Q4: What key survival pattern is visible?\*\*

A4:

- Female survivors are younger on average than male survivors

- Male non-survivors cluster around 20-30 years old

- Children (lower age ranges) had higher survival rates

\*\*Q5: What does the "split=True" parameter do?\*\*

A5: Splits the violin into two halves for direct comparison of survival status within each gender.

\*\*Q6: What does the width of a violin plot represent?\*\*

A6: The density of data points at different ages. Wider sections = more people in that age range.

---

#### \*\*Technical Aspects\*\*

\*\*Q7: Why set `figsize=(8,6)`?\*\*

A7: Adjusts the plot dimensions for better readability (width=8 inches, height=6 inches).

\*\*Q8: How would you customize the legend labels?\*\*

A8: Using `labels=['No', 'Yes']` maps 0/1 survival values to meaningful text.

\*\*Q9: What alternative to `dropna()` could be used?\*\*

A9: Imputation (e.g., filling missing ages with median values) using `fillna()`.

---

#### \*\*Advanced Analysis\*\*

\*\*Q10: What survival advantage is shown for children?\*\*

A10: The lower quartile (boxplot) and density bulge (violinplot) for survivors are skewed toward younger ages, reflecting "women and children first" policy.

\*\*Q11: How would you test if age-survival correlation is statistically significant?\*\*

A11: Use a hypothesis test:

```python

from scipy.stats import ttest\_ind

survived = titanic[titanic['survived']==1]['age']

not\_survived = titanic[titanic['survived']==0]['age']

ttest\_ind(survived, not\_survived)

```

\*\*Q12: Why do male non-survivors cluster around 20-30?\*\*

A12: Likely crew members and third-class passengers (working-age males with lower priority for lifeboats).

---

#### \*\*Syllabus-Related\*\*

\*\*Q13: Name 3 features of a boxplot\*\*

A13:

1. Median line

2. Box (IQR: 25th-75th percentile)

3. Whiskers (1.5×IQR)

4. Outliers (dots beyond whiskers)

\*\*Q14: What is kernel density estimation?\*\*

A14: A smoothed histogram used in violinplots to estimate data distribution.

\*\*Q15: When to use boxplots vs violinplots?\*\*

A15:

- \*\*Boxplots\*\*: Compare summary statistics

- \*\*Violinplots\*\*: Show detailed distribution shapes

---

### \*\*Key Insights\*\*

1. \*\*Gender Bias\*\*: Females had higher survival rates

2. \*\*Age Factor\*\*: Children prioritized in rescues

3. \*\*Crew Demographics\*\*: Many young male non-survivors

This code helps visualize the famous "women and children first" survival pattern from the Titanic disaster. Would you like to explore other survival factors like passenger class? 😊