```
# Import libraries

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns


# Load the dataset

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


# Data Cleaning

## Inspect the data

print(df.head())

print(df.info())

print(df.describe())


## Handle missing values

df['Age'].fillna(df['Age'].median(), inplace=True)

df['Embarked'].fillna(df['Embarked'].mode()[0], inplace=True)

df.drop_duplicates(inplace=True)


## Feature Engineering

df['Title'] = df['Name'].apply(lambda x: x.split(',')[1].split('.')[0].strip())
```

```
df['Sex'] = df['Sex'].map({'male': 0, 'female': 1})

df['Embarked'] = df['Embarked'].map({'C': 0, 'Q': 1, 'S': 2})


# Exploratory Data Analysis (EDA)

## Univariate Analysis

### Age Distribution

plt.figure(figsize=(10, 5))

sns.histplot(df['Age'], bins=30, kde=True)

plt.title('Age Distribution')

plt.xlabel('Age')

plt.ylabel('Frequency')

plt.show()


### Survival Count

plt.figure(figsize=(10, 5))

sns.countplot(x='Survived', data=df)

plt.title('Survival Count')

plt.xlabel('Survived')

plt.ylabel('Count')

plt.show()


## Bivariate Analysis

### Survival by Sex
```

```
plt.figure(figsize=(10, 5))

sns.barplot(x='Sex', y='Survived', data=df)

plt.title('Survival by Sex')

plt.xlabel('Sex')

plt.ylabel('Survival Rate')

plt.xticks(ticks=[0, 1], labels=['Male', 'Female'])

plt.show()


### Survival by Pclass

plt.figure(figsize=(10, 5))

sns.barplot(x='Pclass', y='Survived', data=df)

plt.title('Survival by Pclass')

plt.xlabel('Pclass')

plt.ylabel('Survival Rate')

plt.show()


### Age vs. Survival

plt.figure(figsize=(10, 5))

sns.boxplot(x='Survived', y='Age', data=df)

plt.title('Age vs. Survival')

plt.xlabel('Survived')

plt.ylabel('Age')

plt.show()
```

Correlation Analysis

```
plt.figure(figsize=(10, 7))

corr = df.corr()

sns.heatmap(corr, annot=True, cmap='coolwarm', fmt='.2f')

plt.title('Correlation Matrix')

plt
```

Trend Analysis

```
plt.figure(figsize=(12, 8))

sns.catplot(x='Pclass', hue='Sex', col='Survived', kind='count', data=df)

plt.title('Survival Rate by Class and Gender')

plt.show()
```