In [ ]:

```
import pandas as pd
import numpy as np

import matplotlib.pyplot as plt
import seaborn as sns
```

In [ ]:

```
df_org = pd.read_csv('train.csv')
print(df_org.shape)
```

In [ ]:

```
df_org.head(10)
```

In [ ]:

```
df_org.describe()
```

In [ ]:

```
df_org.groupby('Pclass', as_index=False)['Survived'].mean()
```

In [ ]:

```
sns.barplot('Pclass', 'Survived', data=df_org)
plt.show()
```

In [ ]:

```
sns.countplot(x='Pclass', hue='Survived', data=df_org)
plt.legend(loc='upper left', title='Survived')
plt.show()
```

In [ ]:

```
df_org.groupby('Sex', as_index=False)['Survived'].mean()
```

In [ ]:

```
sns.barplot('Sex', 'Survived', data=df_org)
plt.show()
```

In [ ]:

```
sns.countplot(x='Sex', hue='Survived', data=df_org)
plt.legend(loc='upper right', title='Survived')
plt.show()
```

In [ ]:

```
df_org['Embarked'].describe()
```

In [ ]:

```
df_org['Embarked'].unique()
```

In [ ]:

```
df_org['Embarked'].isnull().describe()
```

In [ ]:

```
df_org['Embarked'] = df_org['Embarked'].fillna('S')

df_org.groupby('Embarked', as_index=False)['Survived'].mean()
```

In [ ]:

```
sns.barplot('Embarked', 'Survived', data=df_org)
plt.show()
```

In [ ]:

```
sns.countplot(x='Embarked', hue='Survived', data=df_org)
plt.legend(loc='upper right', title='Survived')
plt.show()
```

In [ ]:

```
df_org.groupby('SibSp', as_index=False)['Survived'].mean()
```

In [ ]:

```
sns.barplot('SibSp', 'Survived', data=df_org)
plt.show()
```

In [ ]:

```
sns.countplot(x='SibSp', hue='Survived', data=df_org)
plt.legend(loc='upper right', title='Survived')
plt.show()
```

In [ ]:

```
df_org.groupby('Parch', as_index=False)['Survived'].mean()
```

In [ ]:

```
sns.barplot('Parch', 'Survived', data=df_org)
plt.show()
```

In [ ]:

```
sns.countplot(x='Parch', hue='Survived', data=df_org)
plt.legend(loc='upper right')
plt.show()
```

In [ ]:

```
plt.figure(figsize=(15,8))
sns.kdeplot(df_org["Age"][df_org.Survived == 1], color="red", shade=True)
sns.kdeplot(df_org["Age"][df_org.Survived == 0], color="blue", shade=True)
plt.legend(['Survived', 'Died'])
plt.title('Density Plot of Age')
plt.grid()
plt.xticks(np.arange(0, 100, 10))
plt.show()
```

In [ ]:

```
plt.figure(figsize=(15,8))
sns.kdeplot(df_org["Fare"][df_org.Survived == 1], color="red", shade=True)
sns.kdeplot(df_org["Fare"][df_org.Survived == 0], color="blue", shade=True)
plt.legend(['Survived', 'Died'])
plt.title('Density Plot of Fare')
plt.grid()
plt.show()
```

In [ ]:

```
df_org['Fare'].describe()
```