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

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

df.head()
```

{"summary":"{\n \"name\": \"df\",\n \"rows\": 418,\n \"fields\": [\n   {\n   \"column\": \"PassengerId\",\n   \"properties\": {\n   \"dtype\": \"number\",\n   \"std\": 120,\n   \"min\": 892,\n   \"max\": 1309,\n   \"num_unique_values\": 418,\n   \"samples\": [\n   1213,\n   1216,\n   1280\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"Survived\",\n   \"properties\": {\n   \"dtype\": \"number\",\n   \"std\": 0,\n   \"min\": 0,\n   \"max\": 1,\n   \"num_unique_values\": 2,\n   \"samples\": [\n   1,\n   0\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"Pclass\",\n   \"properties\": {\n   \"dtype\": \"number\",\n   \"std\": 0,\n   \"min\": 1,\n   \"max\": 3,\n   \"num_unique_values\": 3,\n   \"samples\": [\n   3,\n   2\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"Name\",\n   \"properties\": {\n   \"dtype\": \"string\",\n   \"num_unique_values\": 418,\n   \"samples\": [\n   \"Krekorian, Mr. Neshan\",\n   \"Kreuchen, Miss. Emilie\"\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"Sex\",\n   \"properties\": {\n   \"dtype\": \"category\",\n   \"num_unique_values\": 2,\n   \"samples\": [\n   \"female\",\n   \"male\"\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"Age\",\n   \"properties\": {\n   \"dtype\": \"number\",\n   \"std\": 14.181209235624422,\n   \"min\": 0.17,\n   \"max\": 76.0,\n   \"num_unique_values\": 79,\n   \"samples\": [\n   10.0,\n   34.5\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"SibSp\",\n   \"properties\": {\n   \"dtype\": \"number\",\n   \"std\": 0,\n   \"min\": 0,\n   \"max\": 8,\n   \"num_unique_values\": 7,\n   \"samples\": [\n   0,\n   1\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"Parch\",\n   \"properties\": {\n   \"dtype\": \"number\",\n   \"std\": 0,\n   \"min\": 0,\n   \"max\": 9,\n   \"num_unique_values\": 8,\n   \"samples\": [\n   1,\n   6\n   ],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n   }\n   },\n   {\n   \"column\": \"Ticket\",\n   \"properties\": {\n   \"dtype\": \"string\",

\"num_unique_values\": 363,\n   \"samples\": [\n
\"2673\",\n   \"W./C. 6607\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"Fare\",\n   \"properties\": {\n
\"dtype\": \"number\",\n   \"std\": 55.907576179973844,\n
\"min\": 0.0,\n   \"max\": 512.3292,\n
\"num_unique_values\": 169,\n   \"samples\": [\n
41.5792,\n   57.75\n   ],\n   \"semantic_type\":
\"\",\n   \"description\": \"\"\n   }\n   },\n   {\n
\"column\": \"Cabin\",\n   \"properties\": {\n   \"dtype\":
\"category\",\n   \"num_unique_values\": 76,\n
\"samples\": [\n   \"A21\",\n   \"E45\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"Embarked\",\n   \"properties\":
{\n   \"dtype\": \"category\",\n   \"num_unique_values\":
3,\n   \"samples\": [\n   \"Q\",\n   \"S\"\n
],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n
}\n   }\n   ]\n}","type":"dataframe","variable_name":"df"}

df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 418 entries, 0 to 417
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   PassengerId  418 non-null    int64
 1   Survived     418 non-null    int64
 2   Pclass       418 non-null    int64
 3   Name         418 non-null    object
 4   Sex          418 non-null    object
 5   Age          332 non-null    float64
 6   SibSp        418 non-null    int64
 7   Parch        418 non-null    int64
 8   Ticket       418 non-null    object
 9   Fare         417 non-null    float64
 10  Cabin        91 non-null     object
 11  Embarked     418 non-null    object
dtypes: float64(2), int64(5), object(5)
memory usage: 39.3+ KB
```

df.describe()

{"summary":"{\n \"name\": \"df\",\n \"rows\": 8,\n \"fields\": [\n
{\n   \"column\": \"PassengerId\",\n   \"properties\": {\n
\"dtype\": \"number\",\n   \"std\": 412.1232851470217,\n
\"min\": 120.81045760473994,\n   \"max\": 1309.0,\n
\"num_unique_values\": 7,\n   \"samples\": [\n   418.0,\n
1100.5,\n   1204.75\n   ],\n   \"semantic_type\":
\"\",\n   \"description\": \"\"\n   }\n   },\n   {\n

\"column\": \"Survived\",\n   \"properties\": {\n   \"dtype\":
\"number\",\n   \"std\": 147.6421942886486,\n   \"min\":
0.0,\n   \"max\": 418.0,\n   \"num_unique_values\": 5,\n
\"samples\": [\n   0.36363636363636365,\n   1.0,\n
0.4816221409322309\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"Pclass\",\n   \"properties\": {\n   \"dtype\": \"number\",\n
\"std\": 147.0758997861715,\n   \"min\": 0.8418375519640503,\n
\"max\": 418.0,\n   \"num_unique_values\": 5,\n
\"samples\": [\n   2.2655502392344498,\n   3.0,\n
0.8418375519640503\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"Age\",\n   \"properties\": {\n   \"dtype\": \"number\",\n
\"std\": 109.15868834351015,\n   \"min\": 0.17,\n   \"max\":
332.0,\n   \"num_unique_values\": 8,\n   \"samples\": [\n
30.272590361445783,\n   27.0,\n   332.0\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"SibSp\",\n   \"properties\": {\
n   \"dtype\": \"number\",\n   \"std\": 147.28745840271156,\
n   \"min\": 0.0,\n   \"max\": 418.0,\n
\"num_unique_values\": 6,\n   \"samples\": [\n   418.0,\n
0.4473684210526316,\n   8.0\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"Parch\",\n   \"properties\": {\
n   \"dtype\": \"number\",\n   \"std\": 147.29373273558997,\
n   \"min\": 0.0,\n   \"max\": 418.0,\n
\"num_unique_values\": 5,\n   \"samples\": [\n
0.3923444976076555,\n   9.0,\n   0.9814288785371691\n
],\n   \"semantic_type\": \"\",\n   \"description\": \"\"\n
}\n   },\n   {\n   \"column\": \"Fare\",\n   \"properties\":
{\n   \"dtype\": \"number\",\n   \"std\":
206.20808392426468,\n   \"min\": 0.0,\n   \"max\":
512.3292,\n   \"num_unique_values\": 8,\n   \"samples\": [\n
35.627188489208635,\n   14.4542,\n   417.0\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   }\n]\n}","type":"dataframe"}

```python
gender = df['Sex'].value_counts()

# Plot the pie chart
gender.plot(kind='pie', autopct='%1.1f%%', startangle=90, figsize=(6,
6))

# Add a title
plt.title('Gender Distribution')

# Show the plot
plt.ylabel('')  # Remove the y-label for a cleaner look
plt.show()
```
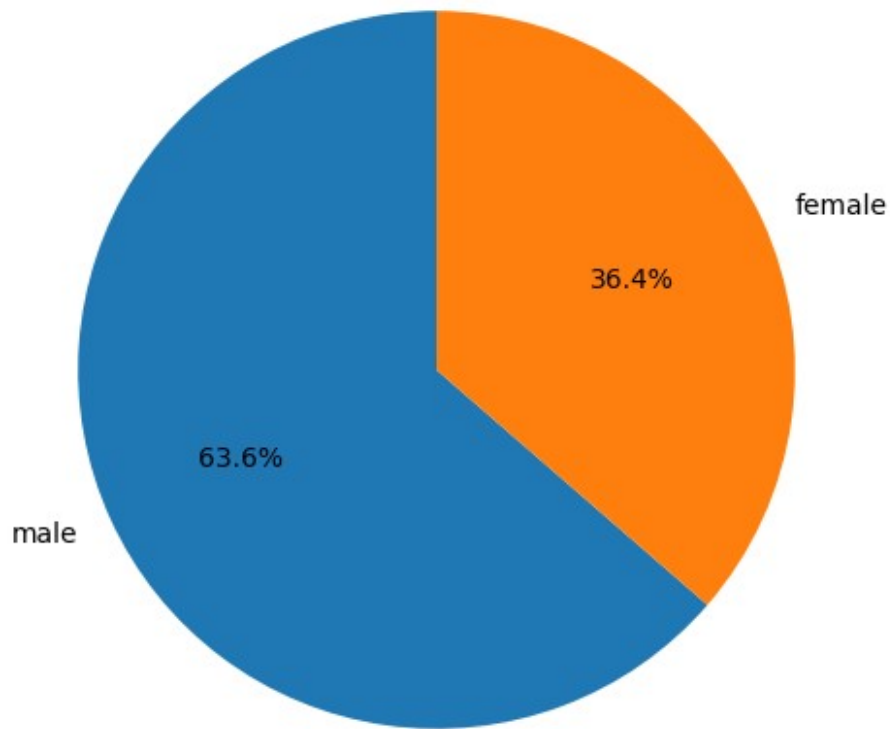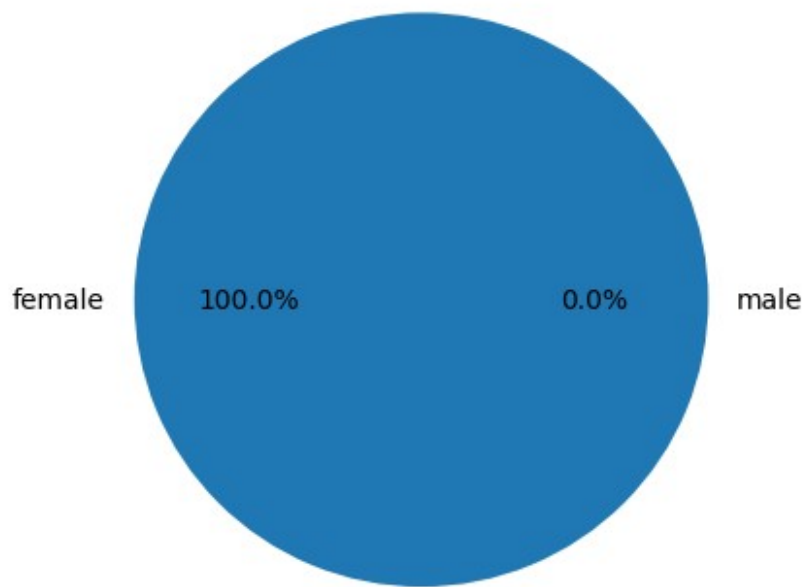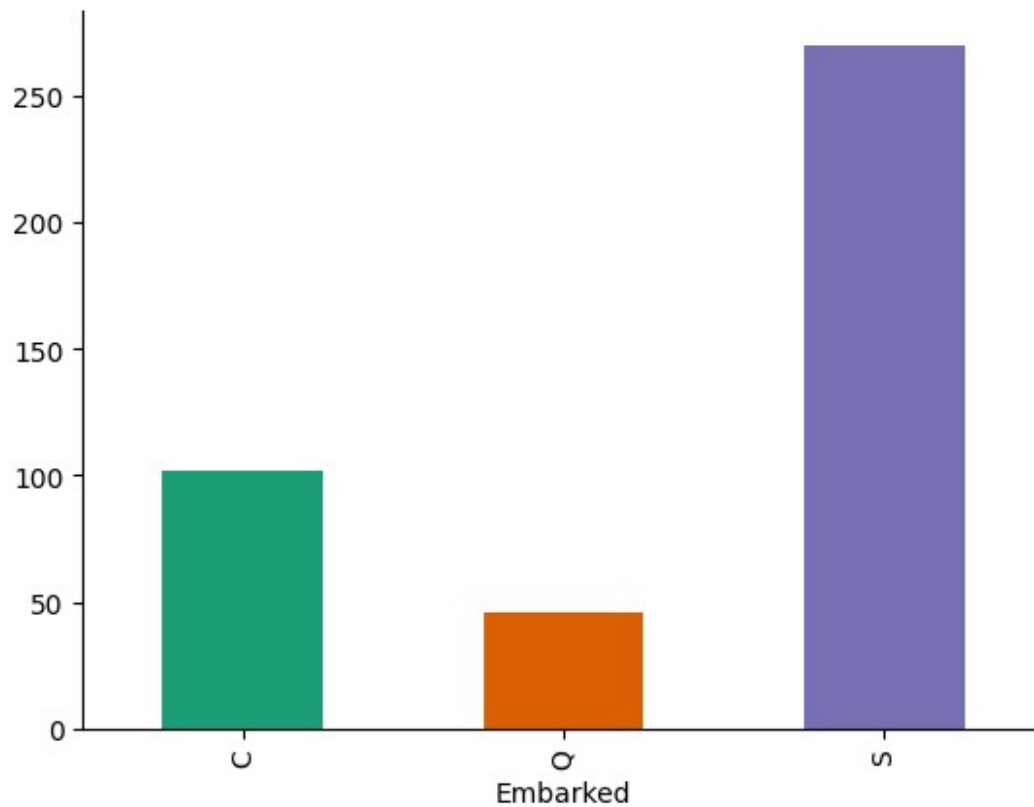
Gender Distribution


```python
survival_rate = df.groupby('Sex')['Survived'].mean()

plt.pie(survival_rate.values, labels=survival_rate.index,
autopct='%1.1f%%')

# Set title
_ = plt.title('Survival Rate by Sex')
```
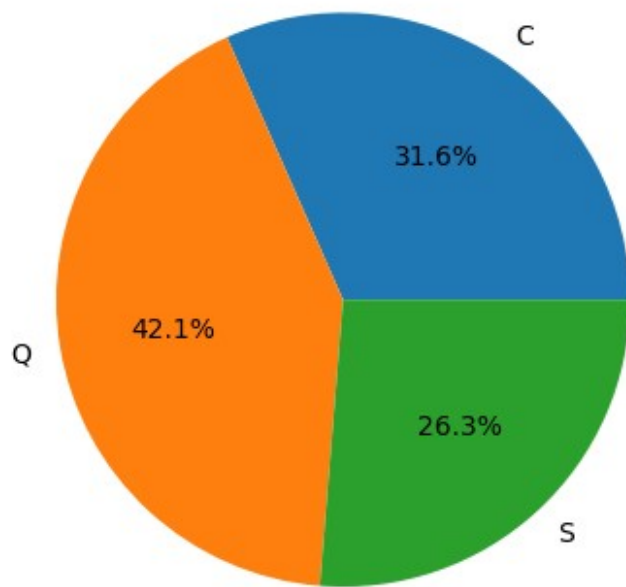
Survival Rate by Sex


```
# @title Embarked

df.groupby('Embarked').size().plot(kind='bar',
color=sns.palettes.mpl_palette('Dark2'))
plt.gca().spines[['top', 'right',]].set_visible(False)
```
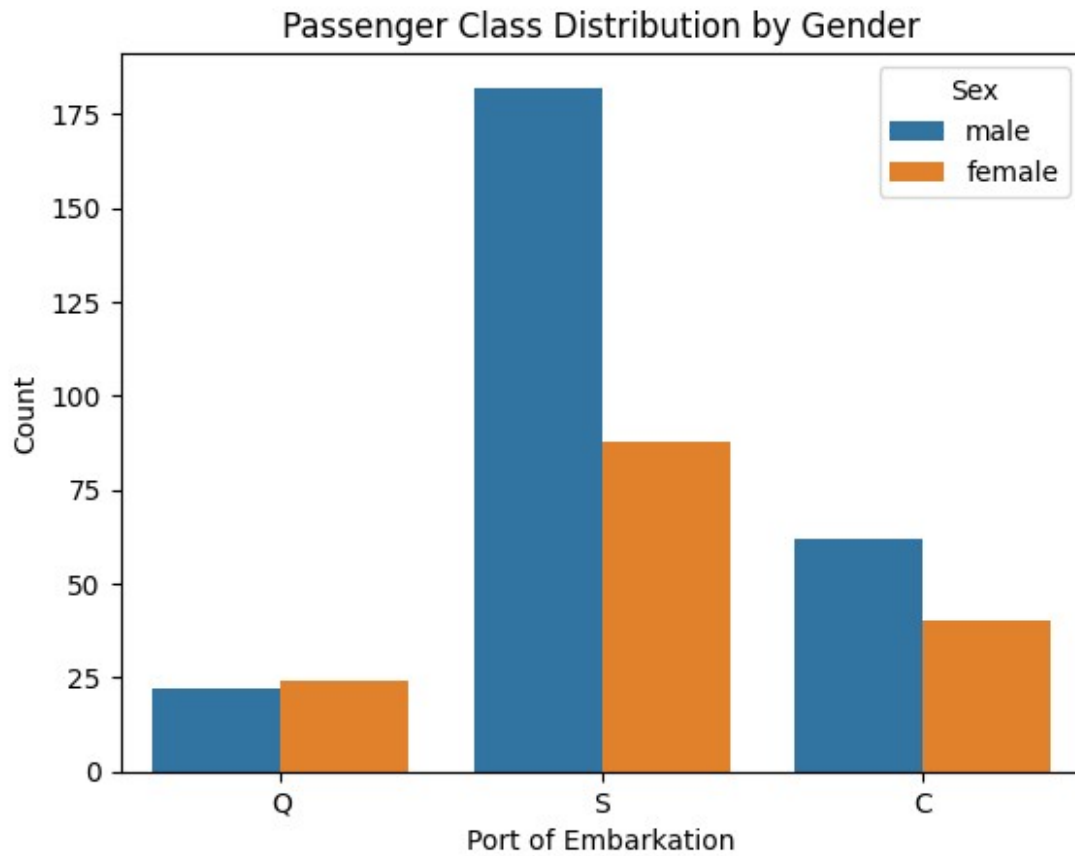
```python
survival_rate = df.groupby('Embarked')['Survived'].mean()

plt.pie(survival_rate.values, labels=survival_rate.index,
autopct='%1.1f%%')

# Set title
_ = plt.title('Survival Rate by Port of Embarkation')
```

Survival Rate by Sex


```
sns.countplot(x='Embarked', hue='Sex', data=df)
plt.xlabel('Port of Embarkation')
plt.ylabel('Count')
plt.title('Passenger Embark Distribution by Gender')
plt.show()
```
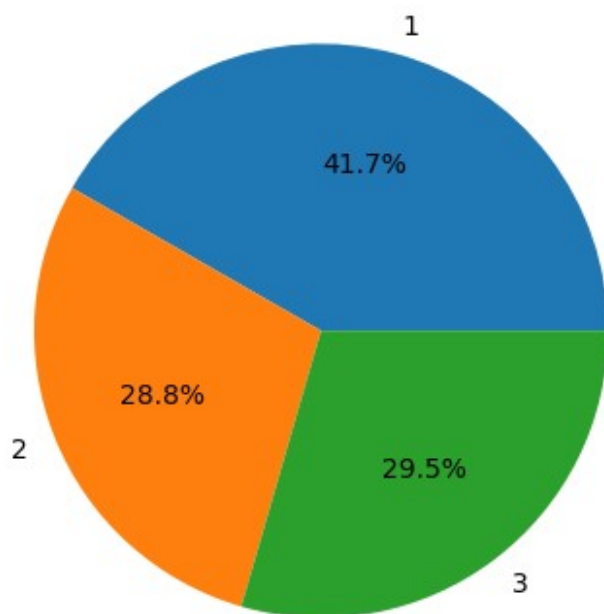
Passenger Class Distribution by Gender

```python
survival_rate = df.groupby('Pclass')['Survived'].mean()

plt.pie(survival_rate.values, labels=survival_rate.index,
autopct='%1.1f%%')

# Set title
_ = plt.title('Survival Rate by Passenger Class')
```

Survival Rate by Passenger Class


```
sns.countplot(x='Pclass', hue='Sex', data=df)

# Add labels and title
plt.xlabel('Passenger Class')
plt.ylabel('Count')
plt.title('Passenger Class Distribution by Gender')

# Show the plot
plt.show()
```

Passenger Class Distribution by Gender