In [19]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:
```
sample=pd.read_csv('train.csv')
```

In [3]:
```
sample.head()
```

Out[3]:

| | PassengerId | Survived | Pclass | Name | Sex | Age | SibSp | Parch | Ticket | Fare | Cabin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 0 | 3 | Braund, Mr. Owen Harris | male | 22.0 | 1 | 0 | A/5 21171 | 7.2500 | NaN |
| **1** | 2 | 1 | 1 | Cumings, Mrs. John Bradley (Florence Briggs Th... | female | 38.0 | 1 | 0 | PC 17599 | 71.2833 | C85 |
| **2** | 3 | 1 | 3 | Heikkinen, Miss. Laina | female | 26.0 | 0 | 0 | STON/O2. 3101282 | 7.9250 | NaN |
| **3** | 4 | 1 | 1 | Futrelle, Mrs. Jacques Heath (Lily May Peel) | female | 35.0 | 1 | 0 | 113803 | 53.1000 | C123 |
| **4** | 5 | 0 | 3 | Allen, Mr. William Henry | male | 35.0 | 0 | 0 | 373450 | 8.0500 | NaN |

In [4]:
```
sample.tail()
```

Out[4]:

| | PassengerId | Survived | Pclass | Name | Sex | Age | SibSp | Parch | Ticket | Fare | Cabin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **886** | 887 | 0 | 2 | Montvila, Rev. Juozas | male | 27.0 | 0 | 0 | 211536 | 13.00 | NaN |
| **887** | 888 | 1 | 1 | Graham, Miss. Margaret Edith | female | 19.0 | 0 | 0 | 112053 | 30.00 | B42 |
| **888** | 889 | 0 | 3 | Johnston, Miss. Catherine Helen "Carrie" | female | NaN | 1 | 2 | W./C. 6607 | 23.45 | NaN |
| **889** | 890 | 1 | 1 | Behr, Mr. Karl Howell | male | 26.0 | 0 | 0 | 111369 | 30.00 | C148 |
| **890** | 891 | 0 | 3 | Dooley, Mr. Patrick | male | 32.0 | 0 | 0 | 370376 | 7.75 | NaN |

In [6]:
```
sample.shape
```

Out[6]:
```
(891, 12)
```

In [8]:
```
sample.isnull()
```

Out[8]:

| | PassengerId | Survived | Pclass | Name | Sex | Age | SibSp | Parch | Ticket | Fare | Cabin | Emb |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | False | False | False | False | False | False | False | False | False | False | True | |
| **1** | False | False | False | False | False | False | False | False | False | False | False | |
| **2** | False | False | False | False | False | False | False | False | False | False | True | |
| **3** | False | False | False | False | False | False | False | False | False | False | False | |
| **4** | False | False | False | False | False | False | False | False | False | False | True | |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| **886** | False | False | False | False | False | False | False | False | False | False | True | |
| **887** | False | False | False | False | False | False | False | False | False | False | False | |
| **888** | False | False | False | False | False | True | False | False | False | False | True | |
| **889** | False | False | False | False | False | False | False | False | False | False | False | |
| **890** | False | False | False | False | False | False | False | False | False | False | True | |

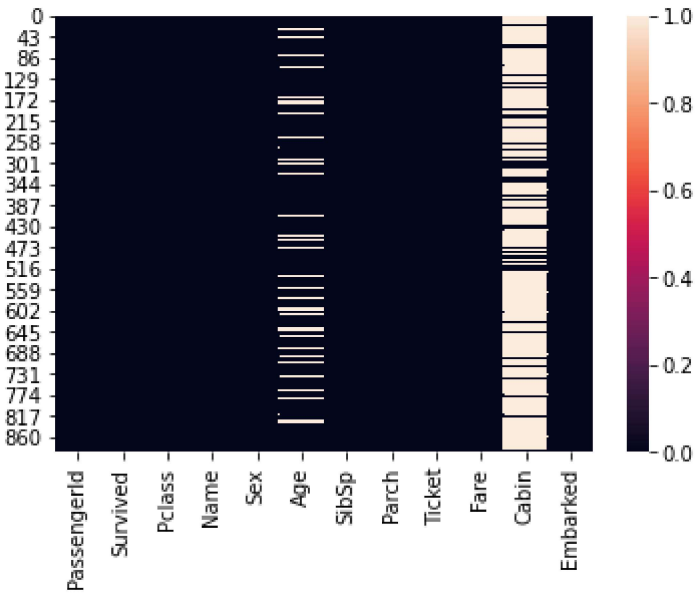
891 rows × 12 columns

In [11]:
```
sample.isnull().sum()
```

Out[11]:
```
PassengerId      0
Survived         0
Pclass           0
Name             0
Sex              0
Age            177
SibSp            0
Parch            0
Ticket           0
Fare             0
Cabin          687
Embarked         2
dtype: int64
```

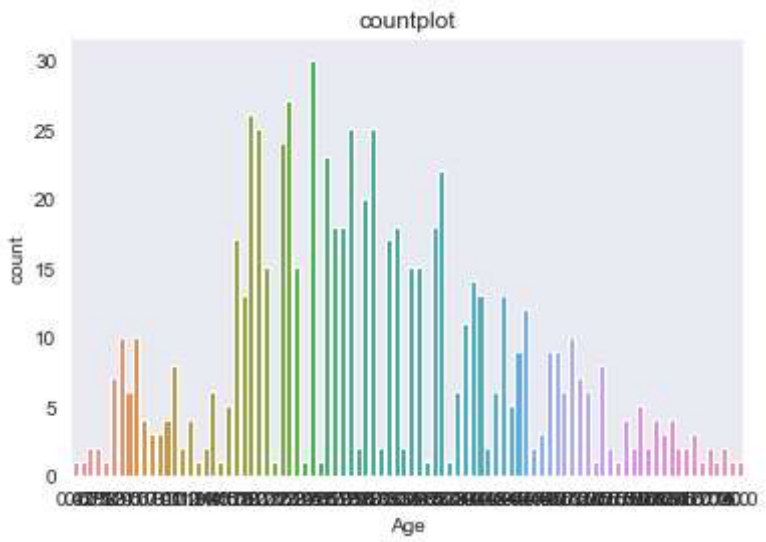
In [20]:
```
sns.heatmap(sample.isnull())
```

Out[20]: <AxesSubplot:>



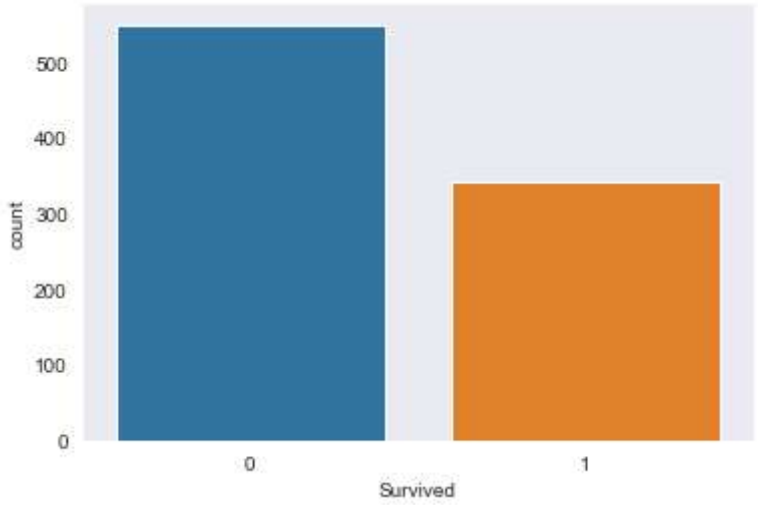
In [27]:
```
sns.set_style('dark')
sns.countplot(x='Age',data=sample)
plt.title('countplot')
```

Out[27]:
```
Text(0.5, 1.0, 'countplot')
```



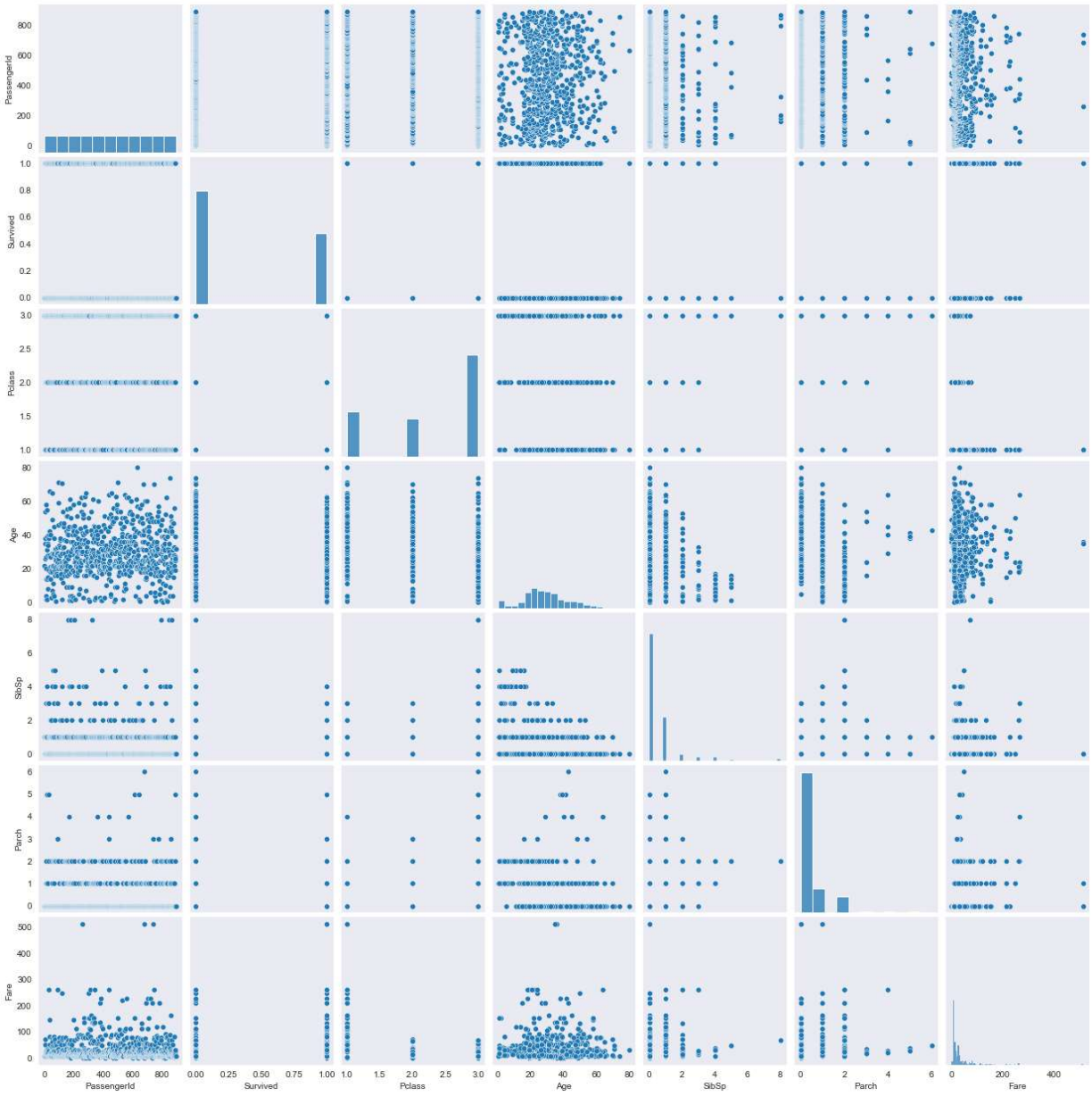

In [28]:
```
sns.countplot(x='Survived',data=sample)
```

Out[28]:
```
<AxesSubplot:xlabel='Survived', ylabel='count'>
```



In [29]:
```
sns.pairplot(sample)
```

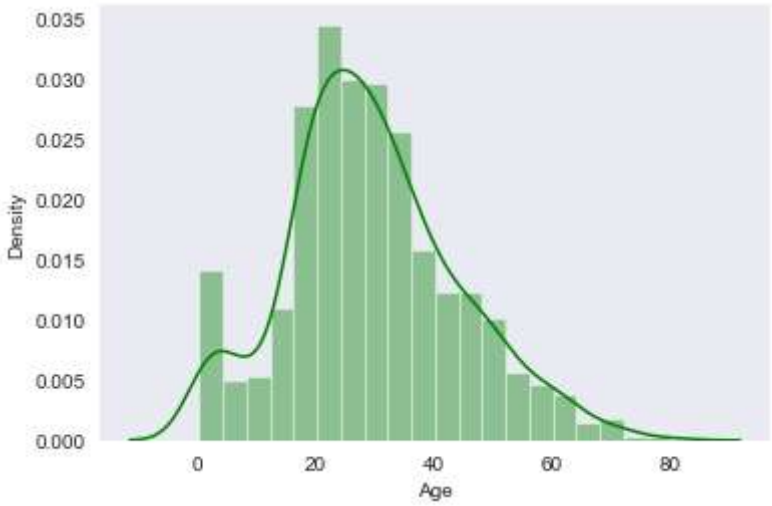
Out[29]:
```
<seaborn.axisgrid.PairGrid at 0x1b63f490b20>
```



In [30]:
```
sns.distplot(sample['Age'].dropna(),color='green',bins=20)
```

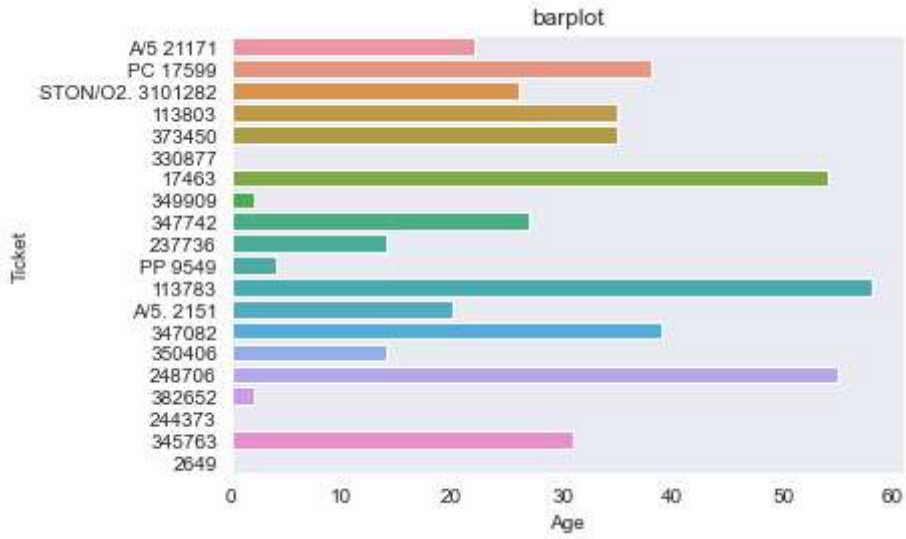
```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Out[30]: <AxesSubplot:xlabel='Age', ylabel='Density'>



In [33]:
```
sns.barplot(x='Age',y='Ticket',data=sample.head(20))
plt.title('barplot')
```

Out[33]: Text(0.5, 1.0, 'barplot')



In [35]:
```
sns.lmplot(x='PassengerId',y='Age',data=sample.head(20),hue='Survived')
plt.title('LMplot')
```

Out[35]: Text(0.5, 1.0, 'LMplot')