

In [1]:

```
import matplotlib.pyplot as plt #grid control and old style graphs
import seaborn as sns #modern library built on top of matplotlib with great collection of graphs
```

In [3]:

```
import pandas as pd
df=pd.read_csv('titanic.csv')
```

In [3]:

df

Out[3]:

	PassengerId	Survived	Pclass	Name	Gender	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S
...	...	...	...	...	...	...	...	...	...	...	...	...
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	NaN	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	B42	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	NaN	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C148	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	NaN	Q

891 rows × 12 columns

In [129]:

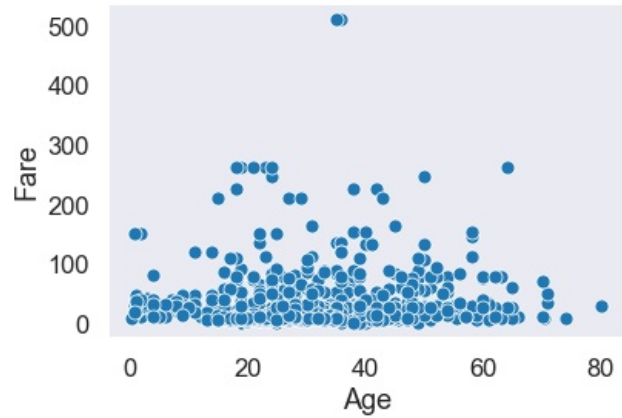
```
sns.set_context('talk') #select from talk poster notebook paper. Sets fonts and settings accordinly
sns.set_style('dark') #choose from white, dark, darkgrid, whitegrid and ticks. Sets background
```

In [8]:

```
sns.scatterplot(x='Age',y='Fare',data=df) #shows relationship between 2 features
```

Out[8]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8eb929790>



In [5]:

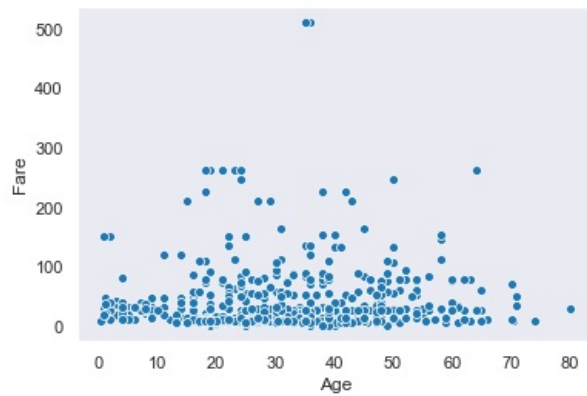
```
sns.set_context('notebook')
sns.set_style('dark')
```

In [10]:

```
sns.scatterplot(x='Age',y='Fare',data=df)
```

Out[10]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8e9101430>

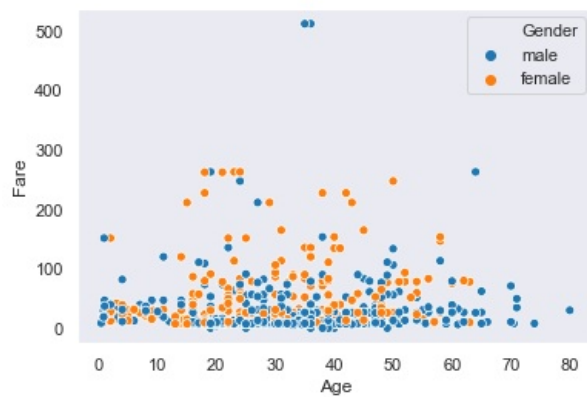


In [12]:

```
sns.scatterplot(x='Age',y='Fare',hue='Gender',data=df) #add a hue. This gives a new dimension to plot.
```

Out[12]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8ebaee760>

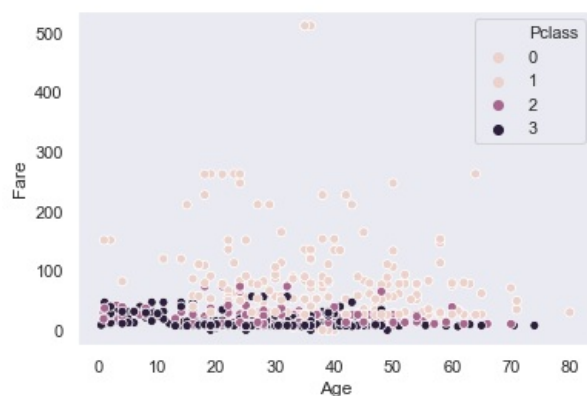


In [13]:

```
sns.scatterplot(x='Age',y='Fare',hue='Pclass',data=df) #hue needs to be categorical
```

Out[13]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8e9089100>

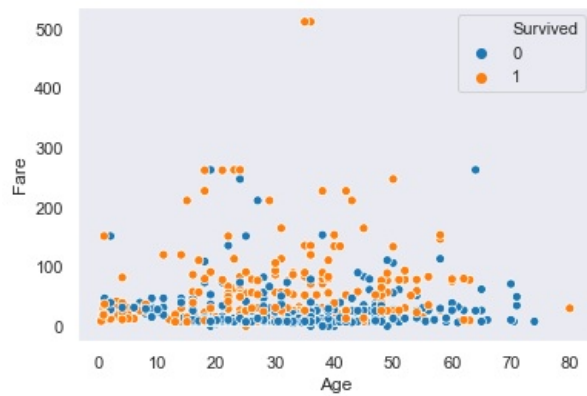


In [14]:

```
sns.scatterplot(x='Age',y='Fare',hue='Survived',data=df) #hue needs to be categorial
```

Out[14]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8e8ffd8b0>

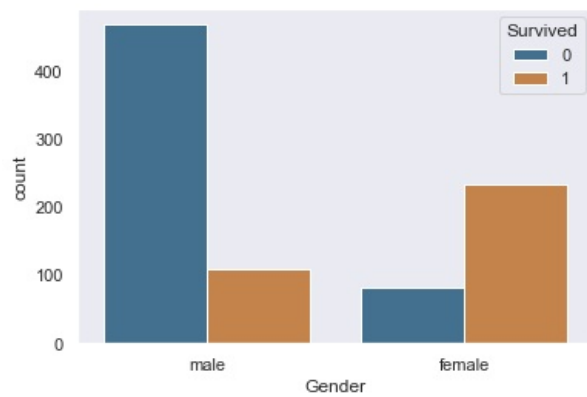


In [23]:

```
sns.countplot(df['Gender'],hue=df['Survived'],saturation=0.5) #saturation controls how bright or dull colours are
```

Out[23]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8e8296760>

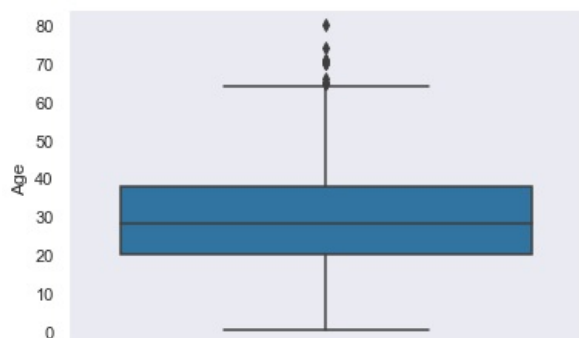


In [24]:

```
sns.boxplot(y=df['Age'])
```

Out[24]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8e8296b20>

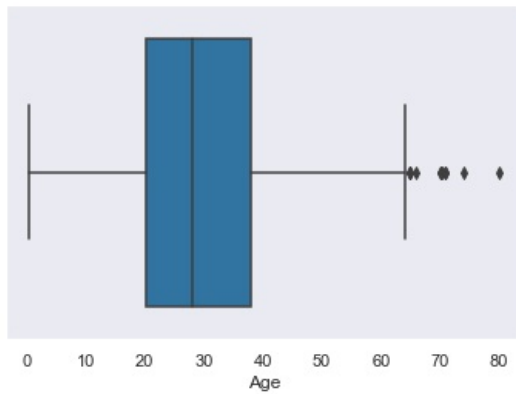


In [25]:

```
sns.boxplot(x=df['Age']) #shows quantile representation
```

Out[25]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8e8122c70>

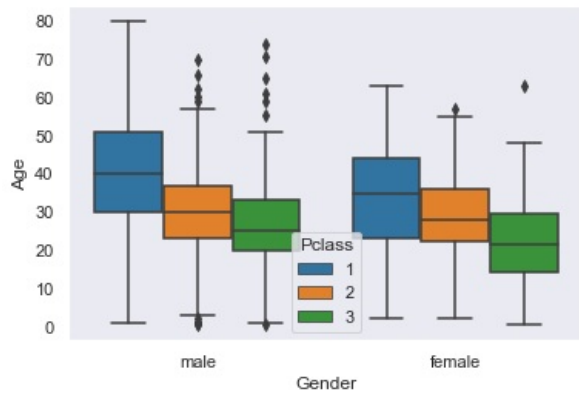


In [35]:

```
sns.boxplot(x=df['Gender'], y=df['Age'], hue=df['Pclass']) #multiple box plots together
```

Out[35]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8e6856f40>

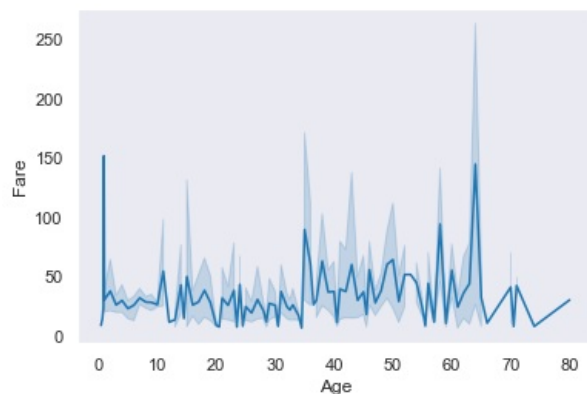


In [31]:

```
sns.lineplot(x=df['Age'], y=df['Fare']) #suitable for continuous type of data
```

Out[31]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8e8c17670>

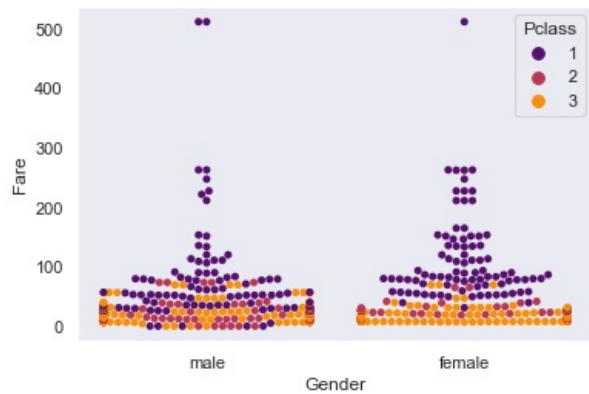


In [42]:

```
sns.swarmplot(x=df['Gender'],y=df['Fare'],hue=df['Pclass'],palette='inferno')
```

Out[42]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8e648bdf0>



In [44]:

```
sns.pairplot(data=df,palette='magma')
```

Out[44]:

<seaborn.axisgrid.PairGrid at 0x7fd8e521f580>



In [47]:

```
# sns.heatmap(df['Age'].reshape(1,))
```

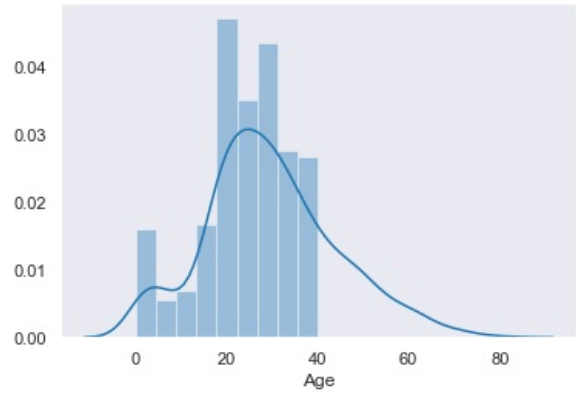
In [57]:

```
import numpy as np

sns.distplot(df['Age'],bins=np.linspace(0,40,10)) #control the range using bin space
```

Out[57]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8db07ca30>



In [58]:

```
df
```

Out[58]:

	PassengerId	Survived	Pclass	Name	Gender	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S
...	...	...	...	...	...	...	...	...	...	...	...	...
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	NaN	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	B42	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	NaN	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C148	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	NaN	Q

891 rows × 12 columns

In [89]:

```
#REPRESENT COUNT OF PASSENGERS SURVIVED IN EVERY EMBARKED DESTINATION CATEGORY HUED BY PCLASS
```

```
temp=df.groupby('Embarked')
temp.groups
```

Out[89]:

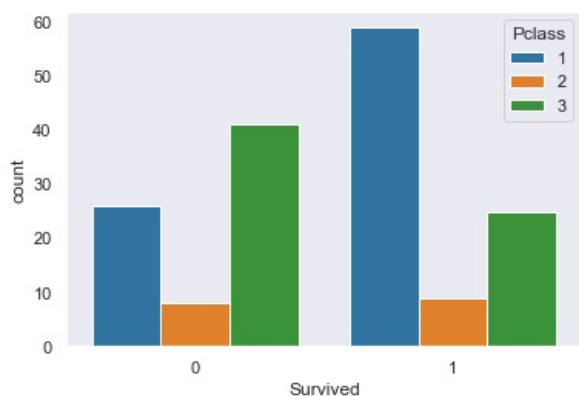
```
{'C': Int64Index([ 1,  9, 19, 26, 30, 31, 34, 36, 39, 42,
                  ...
                  847, 849, 852, 858, 859, 866, 874, 875, 879, 889],
                  dtype='int64', length=168),
 'Q': Int64Index([ 5, 16, 22, 28, 32, 44, 46, 47, 82, 109, 116, 126, 143,
                  156, 171, 186, 188, 196, 198, 208, 214, 241, 245, 260, 264, 274,
                  278, 280, 289, 300, 301, 303, 322, 330, 358, 359, 364, 368, 388,
                  411, 412, 421, 428, 459, 468, 501, 502, 510, 517, 525, 552, 560,
                  573, 593, 612, 613, 626, 629, 653, 654, 657, 680, 697, 703, 718,
                  727, 749, 767, 768, 776, 778, 787, 790, 825, 828, 885, 890],
                  dtype='int64'),
 'S': Int64Index([ 0,  2,  3,  4,  6,  7,  8, 10, 11, 12,
                  ...
                  877, 878, 880, 881, 882, 883, 884, 886, 887, 888],
                  dtype='int64', length=644)}
```

In [97]:

```
df1=temp.get_group('C')[['Survived','Pclass']]
sns.countplot(df1['Survived'],hue=df1['Pclass'])
```

Out[97]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8d8dcb2b0>

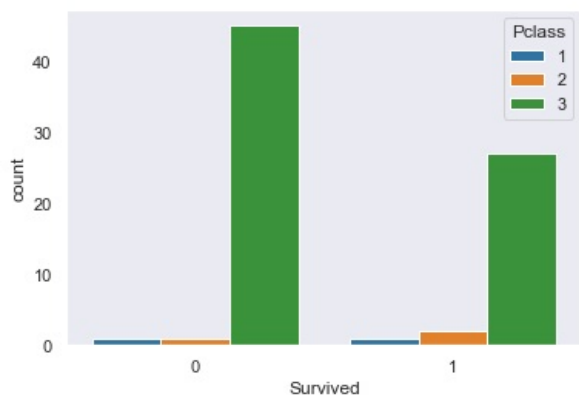


In [96]:

```
df1=temp.get_group('Q')[['Survived','Pclass']]
sns.countplot(df1['Survived'],hue=df1['Pclass'])
```

Out[96]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8d8dcba90>

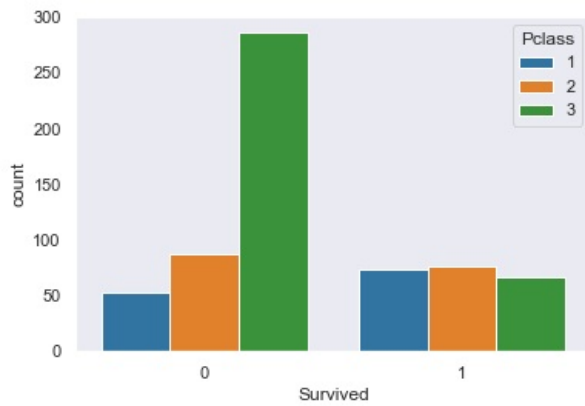


In [95]:

```
df1=temp.get_group('S')[['Survived','Pclass']]
sns.countplot(df1['Survived'],hue=df1['Pclass'])
```

Out[95]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fd8d8432eb0>



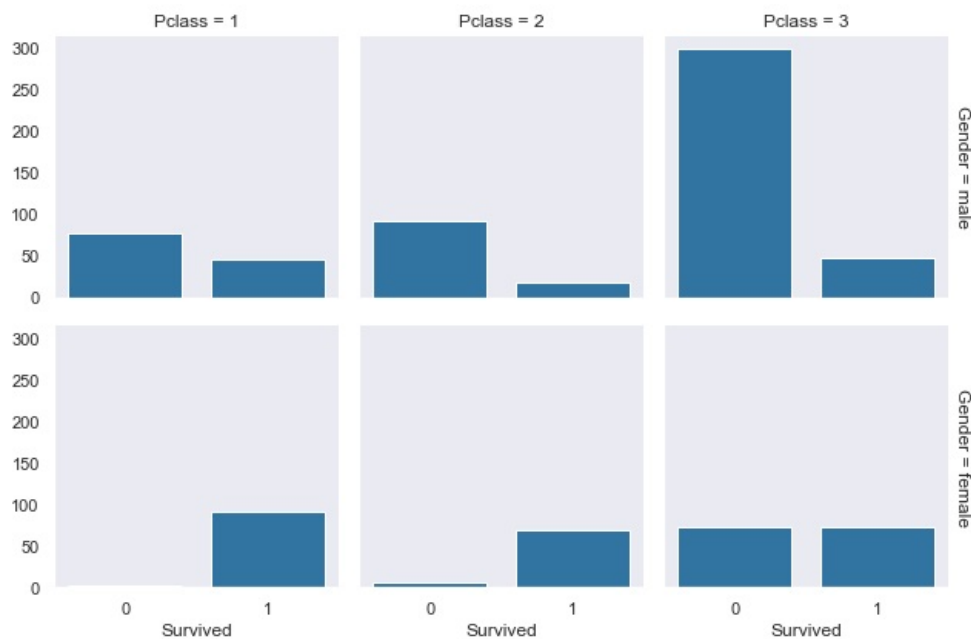
In [101]:

```
grid=sns.FacetGrid(df,row='Gender',col='Pclass',margin_titles=True)
grid.map(sns.countplot,'Survived')
```

/home/harshit/.local/lib/python3.8/site-packages/seaborn/axisgrid.py:723: UserWarning: Using the countplot function without specifying `order` is likely to produce an incorrect plot.  
warnings.warn(warning)

Out[101]:

<seaborn.axisgrid.FacetGrid at 0x7fd8dae6d850>



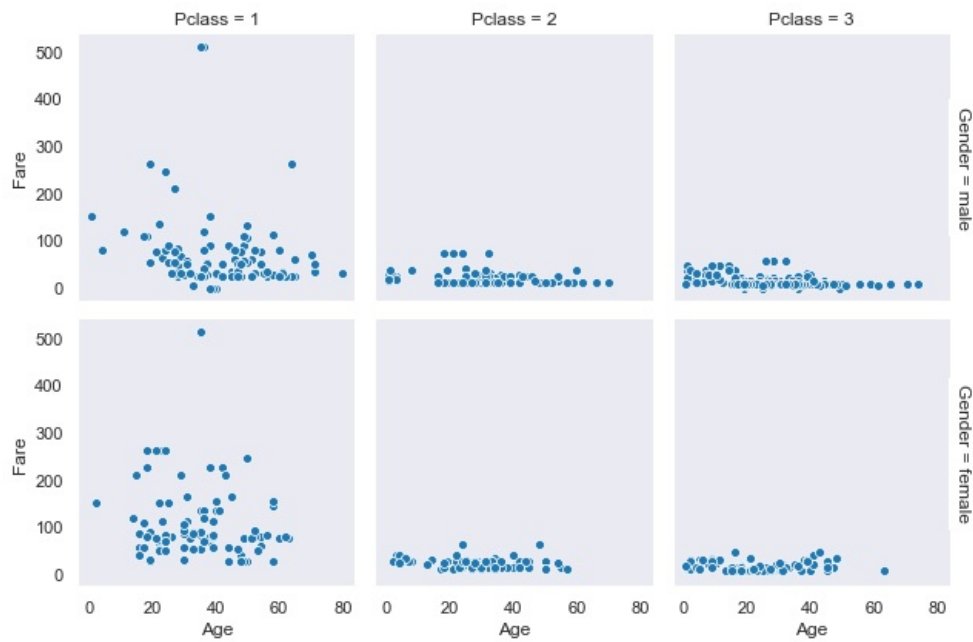


In [98]:

```
grid=sns.FacetGrid(df,row='Gender',col='Pclass',margin_titles=True)
grid.map(sns.scatterplot,'Age','Fare',)
```

Out[98]:

<seaborn.axisgrid.FacetGrid at 0x7fd8d8dcb580>

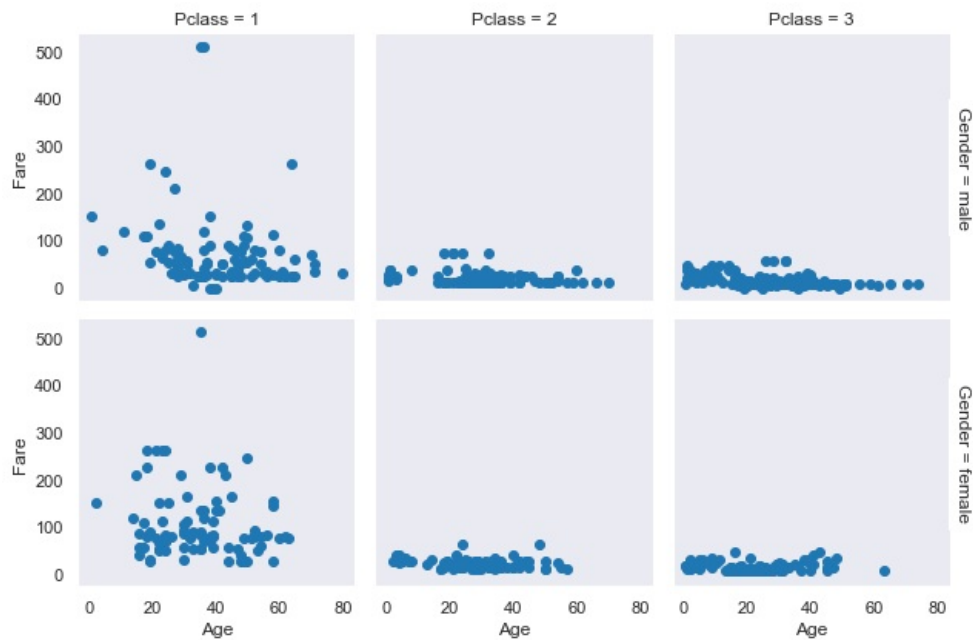


In [100]:

```
grid=sns.FacetGrid(df,row='Gender',col='Pclass',margin_titles=True)
grid.map(plt.scatter,'Age','Fare')
```

Out[100]:

<seaborn.axisgrid.FacetGrid at 0x7fd8d7e094f0>

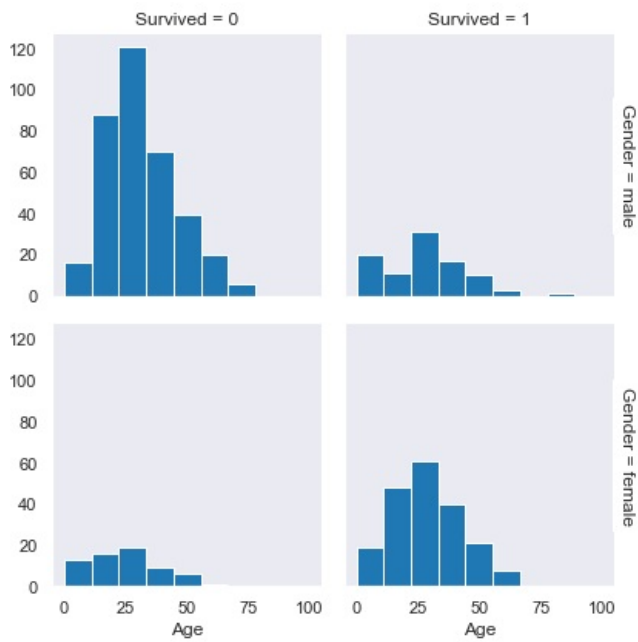


In [65]:

```
grid=sns.FacetGrid(df,row='Gender',col='Survived',margin_titles=True)
grid.map(plt.hist,'Age',bins=np.linspace(0,100,10))
```

Out[65]:

<seaborn.axisgrid.FacetGrid at 0x7fd8da6a2280>

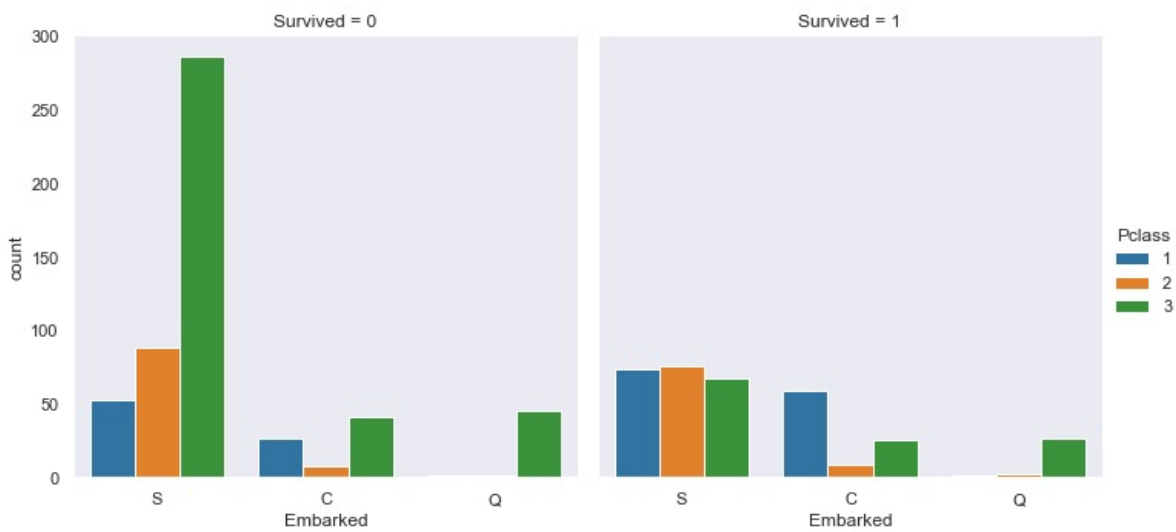


In [102]:

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

Out[102]:

<seaborn.axisgrid.FacetGrid at 0x7fd8d9b14cd0>

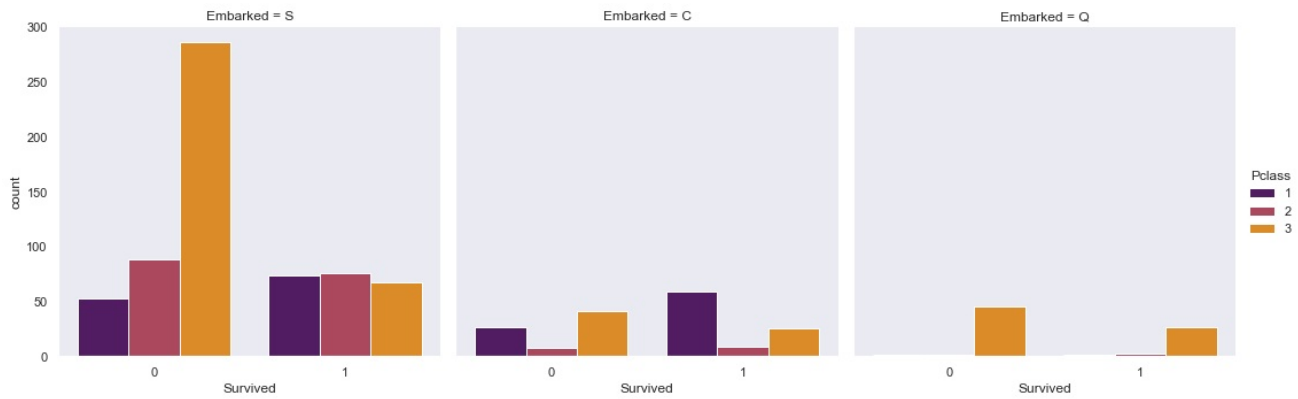


In [107]:

```
sns.catplot(x='Survived',hue='Pclass',col='Embarked',data=df,kind="count",palette='inferno')
```

Out[107]:

<seaborn.axisgrid.FacetGrid at 0x7fd8d738d8b0>

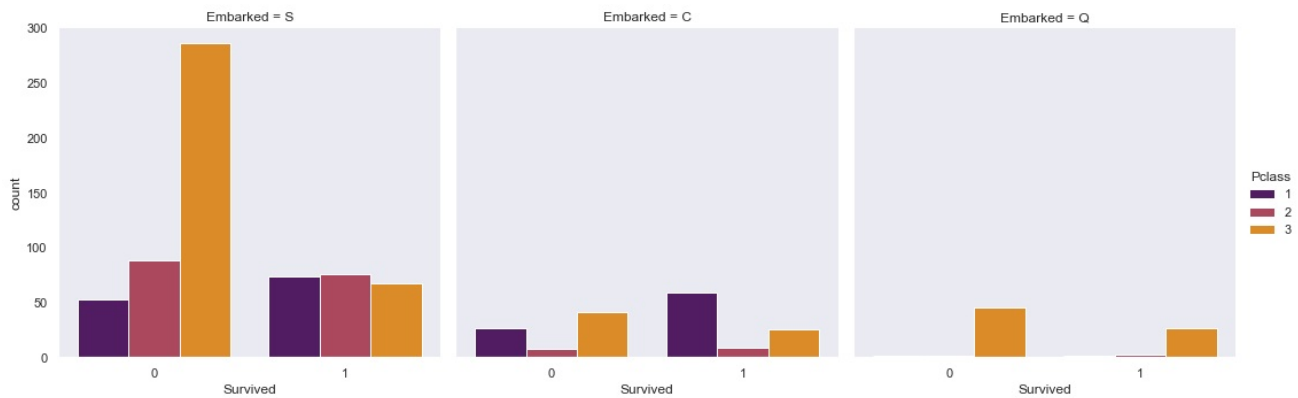


In [112]:

```
sns.catplot(x='Survived',hue='Pclass',col='Embarked',data=df,kind="count",palette='inferno')
```

Out[112]:

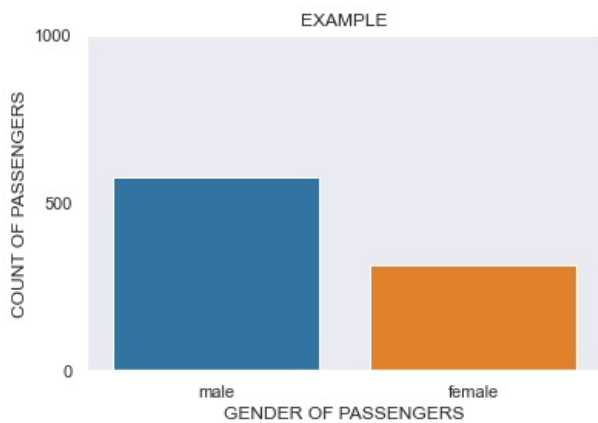
<seaborn.axisgrid.FacetGrid at 0x7fd8dae509d0>



In [118]:

```
g=sns.countplot(df['Gender'])
```

```
plt.xlabel('GENDER OF PASSENGERS')  
plt.ylabel('COUNT OF PASSENGERS')  
plt.yticks([0,500,1000])  
plt.title('EXAMPLE')  
plt.show(g)
```



In [121]:

```
g=sns.countplot(df['Gender'])

plt.xlabel('GENDER OF PASSENGERS')
plt.ylabel('COUNT OF PASSENGERS')
plt.yticks([0,500,1000])
```

8779092028

```
plt.rc('font',size=12)
plt.rc('axes',titlesize=40)
plt.rc('ytick',labelsize=30)
```

```
plt.title('EXAMPLE')
plt.show(g)
```

