

```
In [8]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [36]: data = pd.read_csv('https://raw.githubusercontent.com/dphi-official/Datasets/master/titanic_data.csv')
```

```
In [38]: data
```

Out[38]:

	PassengerId	Survived	Pclass	Name	Sex	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 [40]: data.describe()
```

Out[40]:

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
count	891.000000	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	446.000000	0.383838	2.308642	29.699118	0.523008	0.381594	32.204208
std	257.353842	0.486592	0.836071	14.526497	1.102743	0.806057	49.693429
min	1.000000	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	223.500000	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	446.000000	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	668.500000	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

```
In [42]: data.describe(include = 'object')
```

Out[42]:

	Name	Sex	Ticket	Cabin	Embarked
count	891	891	891	204	889
unique	891	2	681	147	3
top	Braund, Mr. Owen Harris	male	347082	B96 B98	S
freq	1	577	7	4	644

```
In [46]: data[['Pclass', 'Survived']].groupby(['Pclass'], as_index=False).mean().sort_values(by='Survived', ascending
```

Out[46]:

	Pclass	Survived
0	1	0.629630
1	2	0.472826
2	3	0.242363

```
In [48]: data[["Sex", "Survived"]].groupby(['Sex'], as_index=False).mean().sort_values(by='Survived', ascending=False)
```

```
Out[48]:
```

	Sex	Survived
0	female	0.742038
1	male	0.188908

```
In [51]: data[["SibSp", "Survived"]].groupby(['SibSp'], as_index=False).mean().sort_values(by='Survived', ascending=False)
```

```
Out[51]:
```

	SibSp	Survived
1	1	0.535885
2	2	0.464286
0	0	0.345395
3	3	0.250000
4	4	0.166667
5	5	0.000000
6	8	0.000000

```
In [53]: data[["Parch", "Survived"]].groupby(['Parch'], as_index=False).mean().sort_values(by='Survived', ascending=False)
```

```
Out[53]:
```

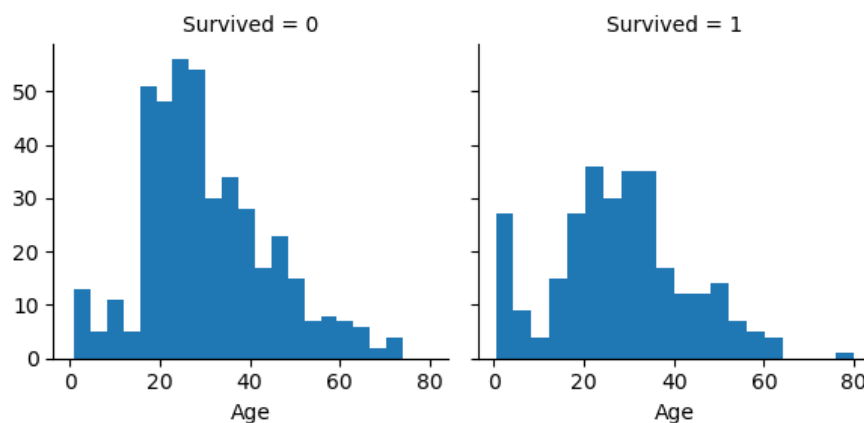
	Parch	Survived
3	3	0.600000
1	1	0.550847
2	2	0.500000
0	0	0.343658
5	5	0.200000
4	4	0.000000
6	6	0.000000

```
In [56]: g = sns.FacetGrid(data, col='Survived')
g.map(plt.hist, 'Age', bins=20)
```

C:\Users\System21\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight

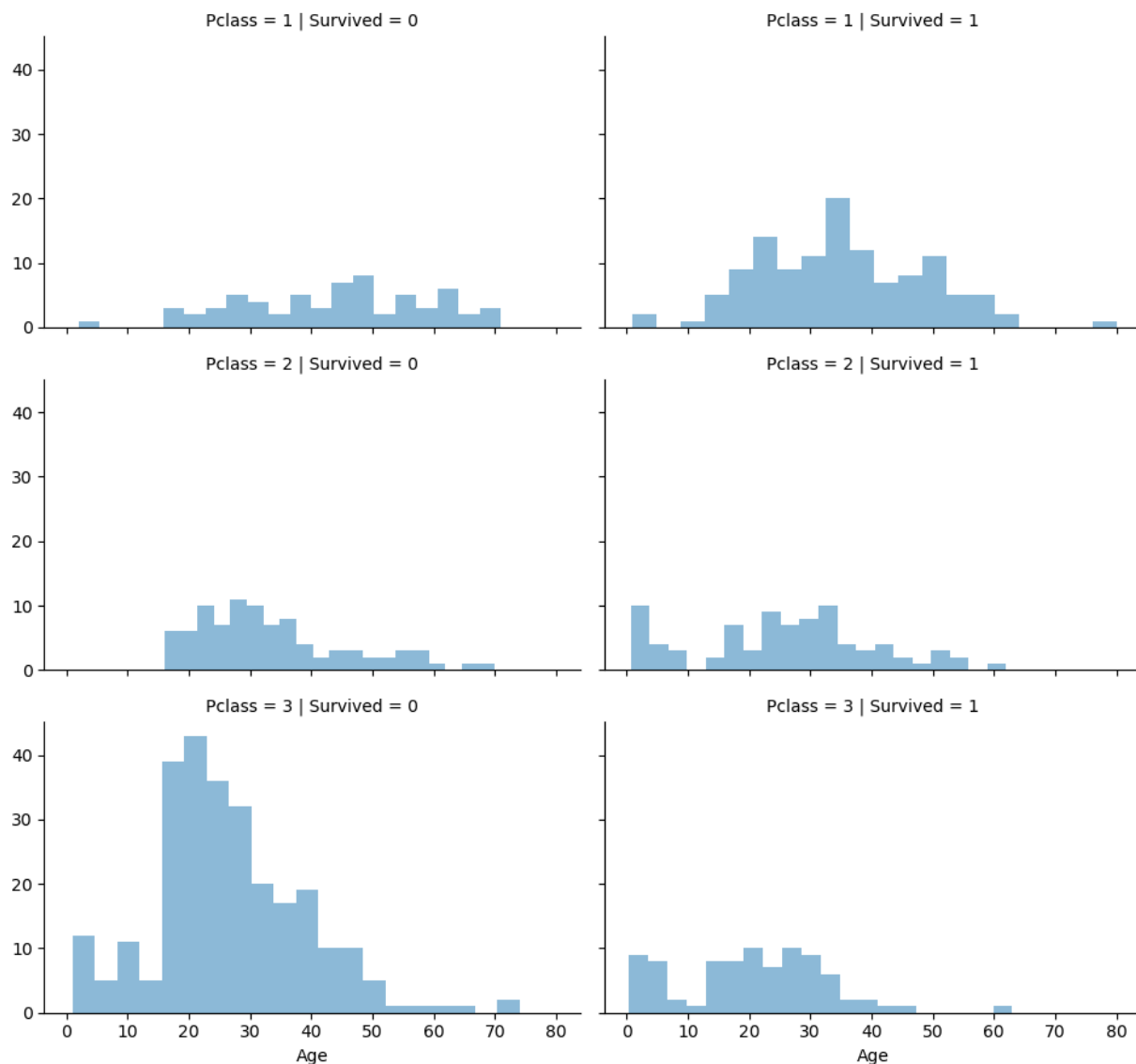
```
self._figure.tight_layout(*args, **kwargs)
```

```
Out[56]: <seaborn.axisgrid.FacetGrid at 0x22949b269d0>
```



```
In [59]: grid = sns.FacetGrid(data, col='Survived', row='Pclass', aspect=1.6)
grid.map(plt.hist, 'Age', alpha=.5, bins=20)
grid.add_legend();
```

C:\Users\System21\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight
self._figure.tight_layout(*args, **kwargs)



```
In [61]: grid = sns.FacetGrid(data, row='Embarked', col='Survived', aspect=1.6)
grid.map(sns.barplot, 'Sex', 'Age', alpha=.5, ci=None)
grid.add_legend()
```

C:\Users\System21\anaconda3\Lib\site-packages\seaborn\axisgrid.py:712: UserWarning: Using the barplot function without specifying `order` is likely to produce an incorrect plot.

warnings.warn(warning)

C:\Users\System21\anaconda3\Lib\site-packages\seaborn\axisgrid.py:848: FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

func(*plot_args, **plot_kwargs)

C:\Users\System21\anaconda3\Lib\site-packages\seaborn\axisgrid.py:848: FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

func(*plot_args, **plot_kwargs)

C:\Users\System21\anaconda3\Lib\site-packages\seaborn\axisgrid.py:848: FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

func(*plot_args, **plot_kwargs)

C:\Users\System21\anaconda3\Lib\site-packages\seaborn\axisgrid.py:848: FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

func(*plot_args, **plot_kwargs)

C:\Users\System21\anaconda3\Lib\site-packages\seaborn\axisgrid.py:848: FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

func(*plot_args, **plot_kwargs)

C:\Users\System21\anaconda3\Lib\site-packages\seaborn\axisgrid.py:848: FutureWarning:

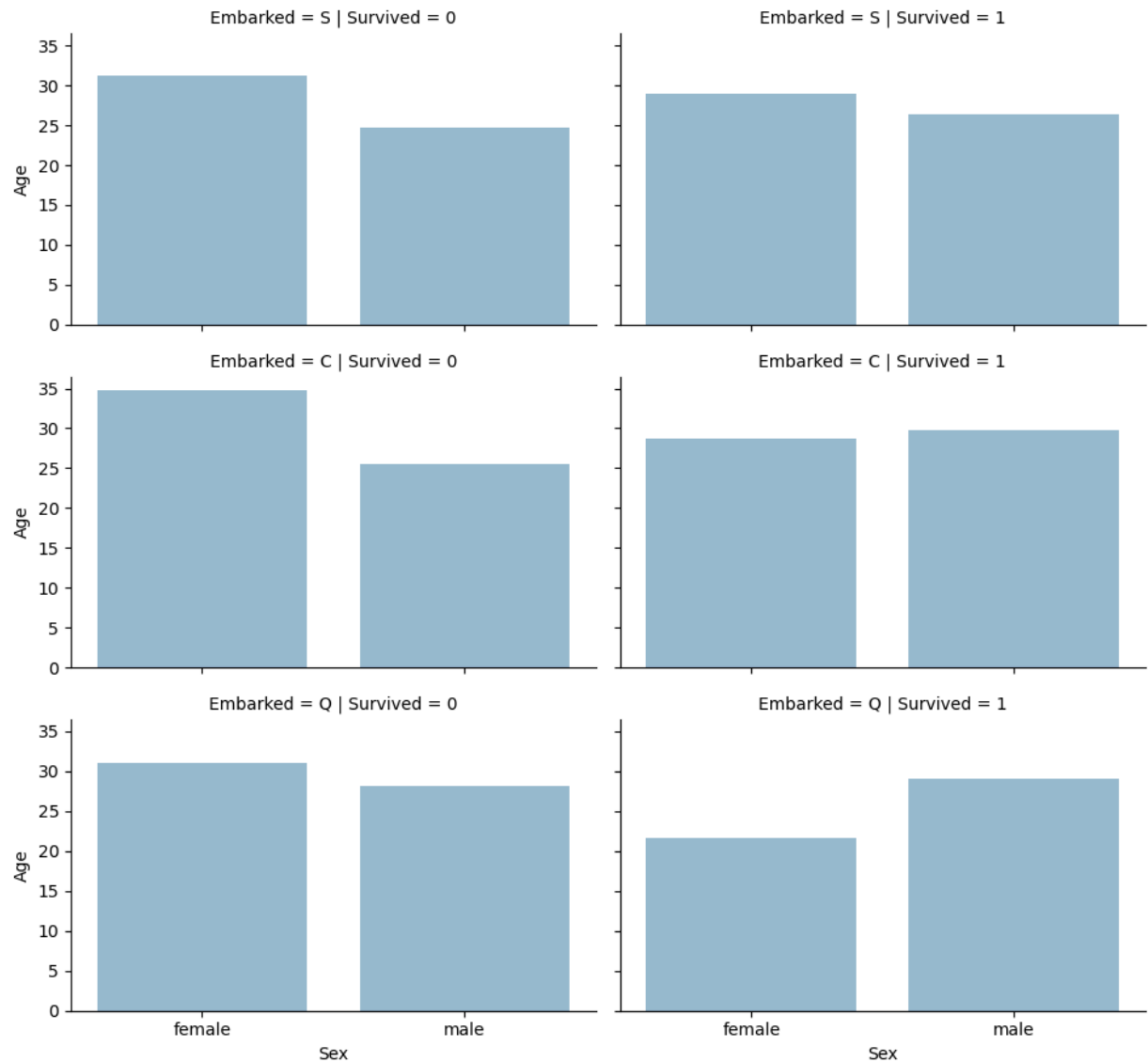
The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

func(*plot_args, **plot_kwargs)

C:\Users\System21\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight

self._figure.tight_layout(*args, **kwargs)

```
Out[61]: <seaborn.axisgrid.FacetGrid at 0x2294bfcbed0>
```

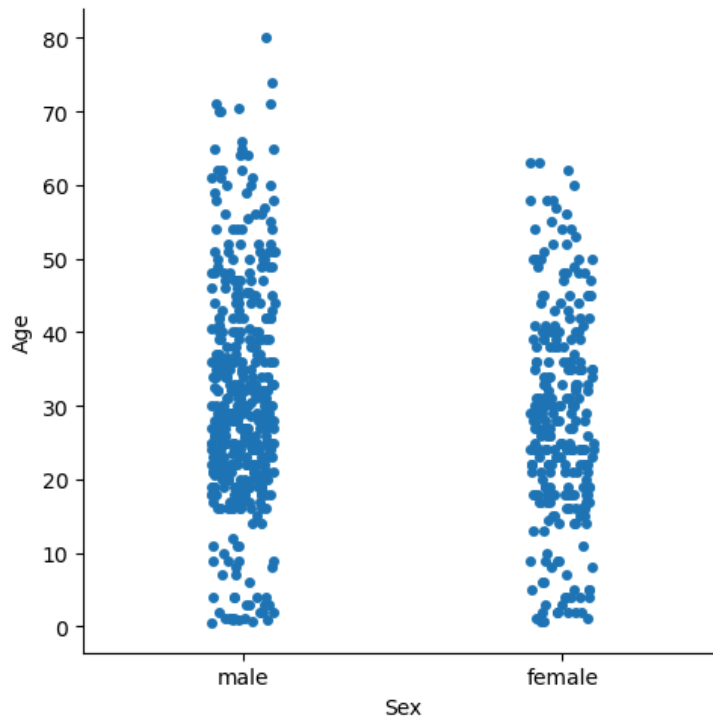


```
In [63]: sns.catplot(x= 'Sex', y = 'Age', data=data, kind = 'strip')
```

C:\Users\System21\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight

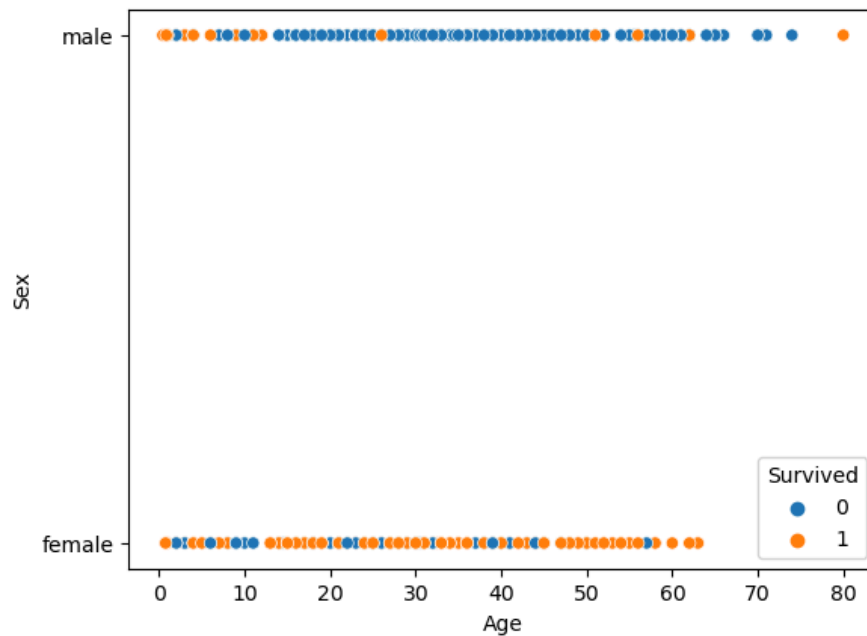
```
self._figure.tight_layout(*args, **kwargs)
```

```
Out[63]: <seaborn.axisgrid.FacetGrid at 0x2294a467d50>
```



```
In [65]: sns.scatterplot(x = 'Age', y = 'Sex', hue = 'Survived', data = data)
```

```
Out[65]: <Axes: xlabel='Age', ylabel='Sex'>
```



```
In [67]: sns.distplot(data['Age'])
```

C:\Users\System21\AppData\Local\Temp\ipykernel_7156\1298243121.py:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

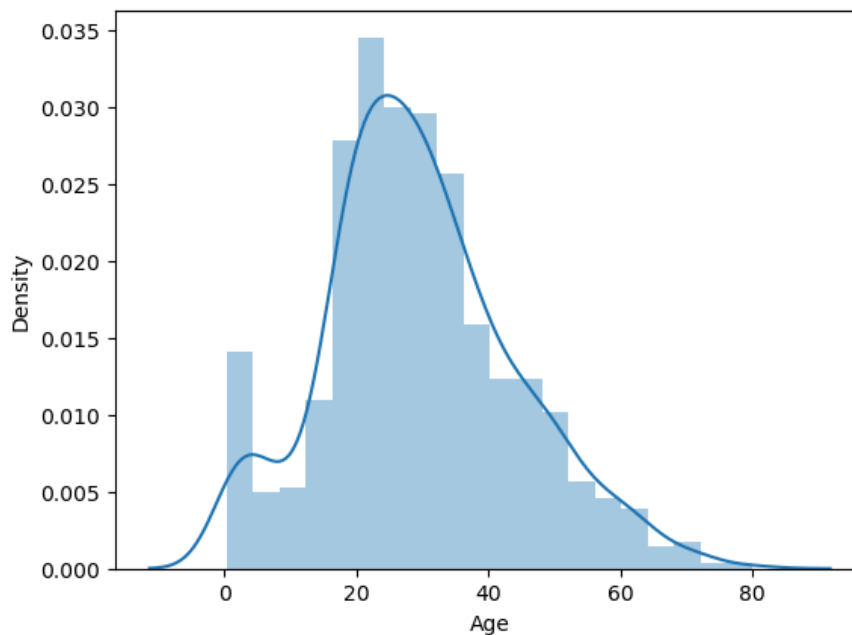
Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see

<https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751> (<https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751>)

```
sns.distplot(data['Age'])
```

Out[67]: <Axes: xlabel='Age', ylabel='Density'>

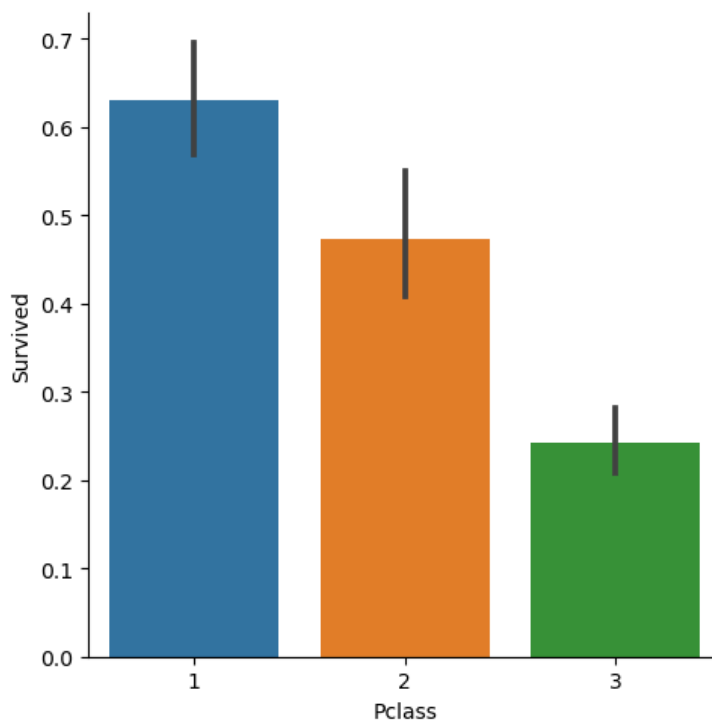


```
In [69]: sns.catplot(x='Pclass', y='Survived', data=data, kind='bar')
```

C:\Users\System21\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has changed to tight

```
self._figure.tight_layout(*args, **kwargs)
```

Out[69]: <seaborn.axisgrid.FacetGrid at 0x22949cce0d0>



```
In [71]: data[['Pclass', 'Survived']].groupby(['Pclass'], as_index=False).mean()
```

Out[71]:

	Pclass	Survived
0	1	0.629630
1	2	0.472826
2	3	0.242363

```
In [73]: data[['Age', 'Survived']].groupby(['Age'], as_index=False).mean()
```

Out[73]:

	Age	Survived
0	0.42	1.0
1	0.67	1.0
2	0.75	1.0
3	0.83	1.0
4	0.92	1.0
...
83	70.00	0.0
84	70.50	0.0
85	71.00	0.0
86	74.00	0.0
87	80.00	1.0

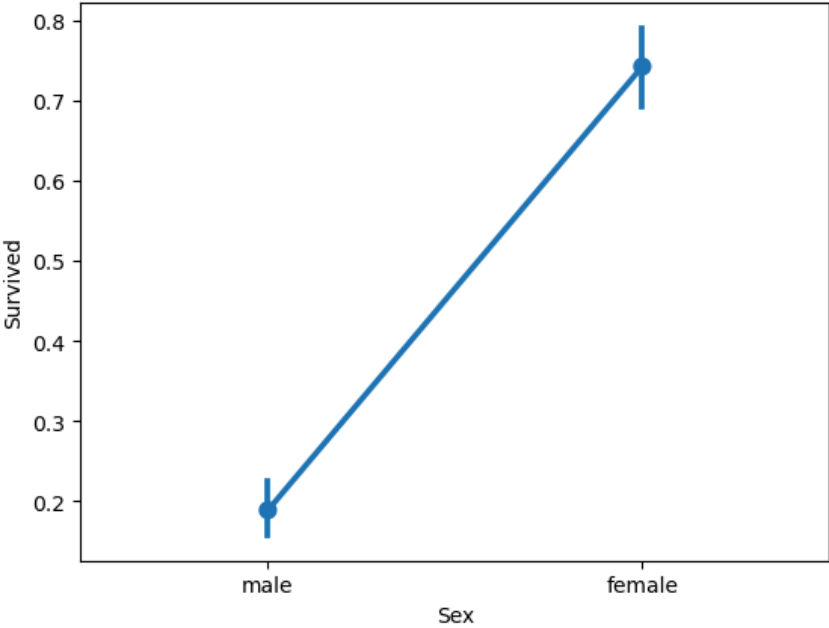
88 rows × 2 columns

```
In [75]: data[['Sex', 'Survived']].groupby(['Sex'], as_index=False).mean()
```

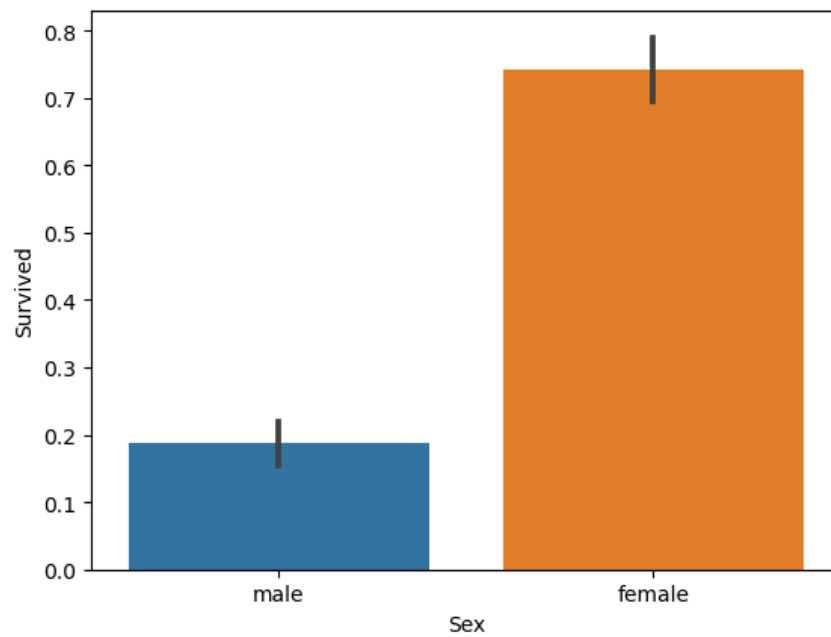
Out[75]:

	Sex	Survived
0	female	0.742038
1	male	0.188908

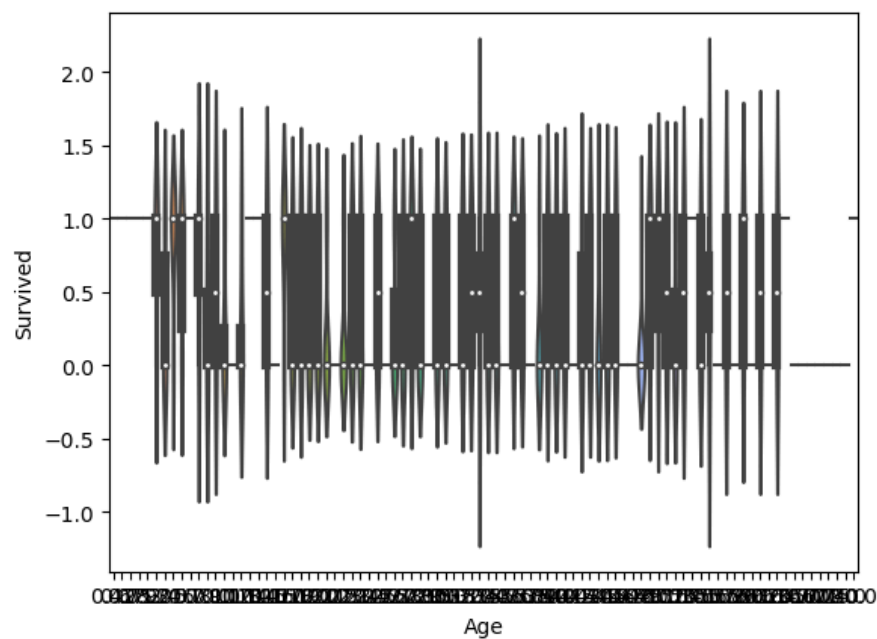
```
In [77]: sns.pointplot(x = "Sex", y = "Survived", data = data)
plt.show()
```




```
In [79]: sns.barplot(x = "Sex", y = "Survived", data = data)
plt.show()
```



```
In [81]: sns.violinplot(x = "Age", y = "Survived", data=data)
plt.show()
```



```
In [82]: sns.swarmplot(x = "Pclass", y ="Survived", data = data)  
plt.show()
```

C:\Users\System21\anaconda3\Lib\site-packages\seaborn\categorical.py:3544: UserWarning: 76.9% of the points cannot be placed; you may want to decrease the size of the markers or use stripplot.
warnings.warn(msg, UserWarning)
C:\Users\System21\anaconda3\Lib\site-packages\seaborn\categorical.py:3544: UserWarning: 72.8% of the points cannot be placed; you may want to decrease the size of the markers or use stripplot.
warnings.warn(msg, UserWarning)
C:\Users\System21\anaconda3\Lib\site-packages\seaborn\categorical.py:3544: UserWarning: 89.8% of the points cannot be placed; you may want to decrease the size of the markers or use stripplot.
warnings.warn(msg, UserWarning)
C:\Users\System21\anaconda3\Lib\site-packages\seaborn\categorical.py:3544: UserWarning: 82.4% of the points cannot be placed; you may want to decrease the size of the markers or use stripplot.
warnings.warn(msg, UserWarning)
C:\Users\System21\anaconda3\Lib\site-packages\seaborn\categorical.py:3544: UserWarning: 79.3% of the points cannot be placed; you may want to decrease the size of the markers or use stripplot.
warnings.warn(msg, UserWarning)
C:\Users\System21\anaconda3\Lib\site-packages\seaborn\categorical.py:3544: UserWarning: 92.3% of the points cannot be placed; you may want to decrease the size of the markers or use stripplot.
warnings.warn(msg, UserWarning)

