In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
data = pd.read_csv('titanic.csv')
data
```

Out[1]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fa
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.25(
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.28:
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.92:
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.10(
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.05(
...	...	...	...	...	...	...	...	...	...	
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.00(
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.00(
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.45(
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.00(
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.75(

891 rows × 12 columns

In [2]:
```
data.shape
```

Out[2]:
```
(891, 12)
```

In [3]:
```
data.describe()
```

Out[3]:

	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 [4]:
```
data.describe(include = 'object')
```

Out[4]:

	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 [5]:
```
data.isnull().sum()
```

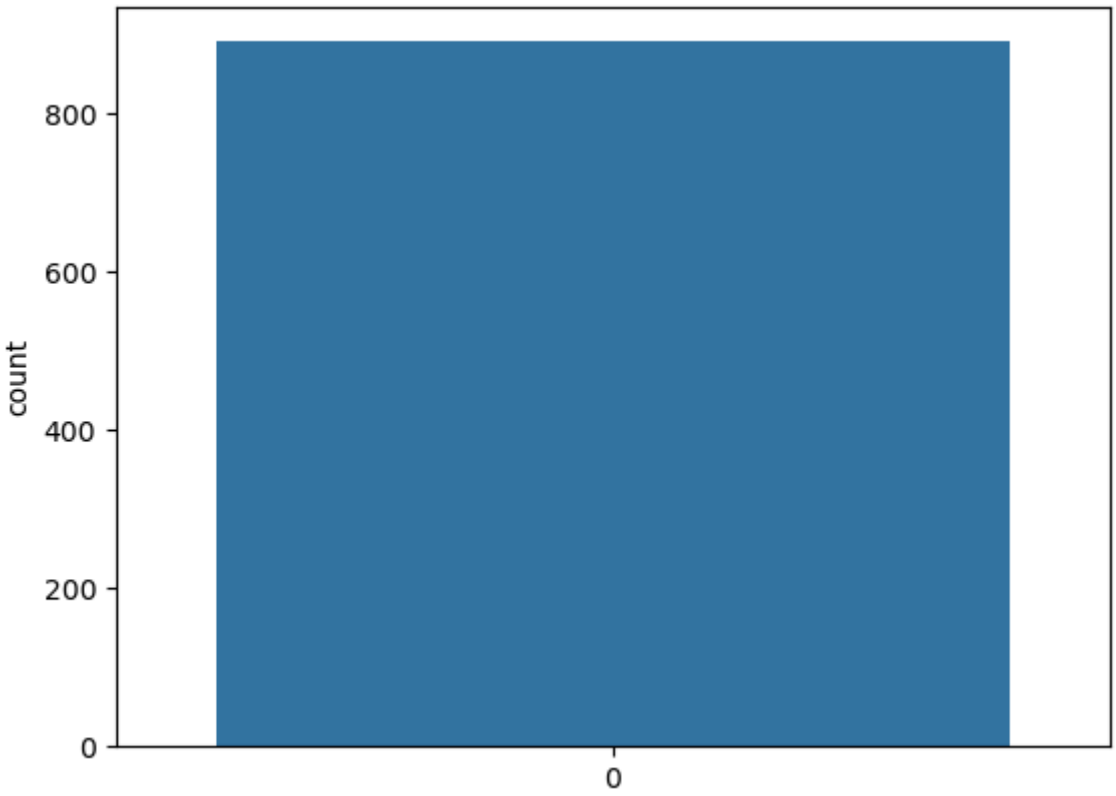
Out[5]:
```
PassengerId      0
Survived         0
Pclass           0
Name             0
Sex              0
Age            177
SibSp            0
Parch            0
Ticket           0
Fare             0
Cabin          687
Embarked         2
dtype: int64
```

In [6]:
```
data['Age'] = data['Age'].fillna(np.mean(data['Age']))
data['Cabin'] = data['Cabin'].fillna(data['Cabin'].mode()[0])
data['Embarked'] = data['Embarked'].fillna(data['Embarked'].mode()[0])
data.isnull().sum()
```

Out[6]:
```
PassengerId    0
Survived       0
Pclass         0
Name           0
Sex            0
Age            0
SibSp          0
Parch          0
Ticket         0
Fare           0
Cabin          0
Embarked       0
dtype: int64
```

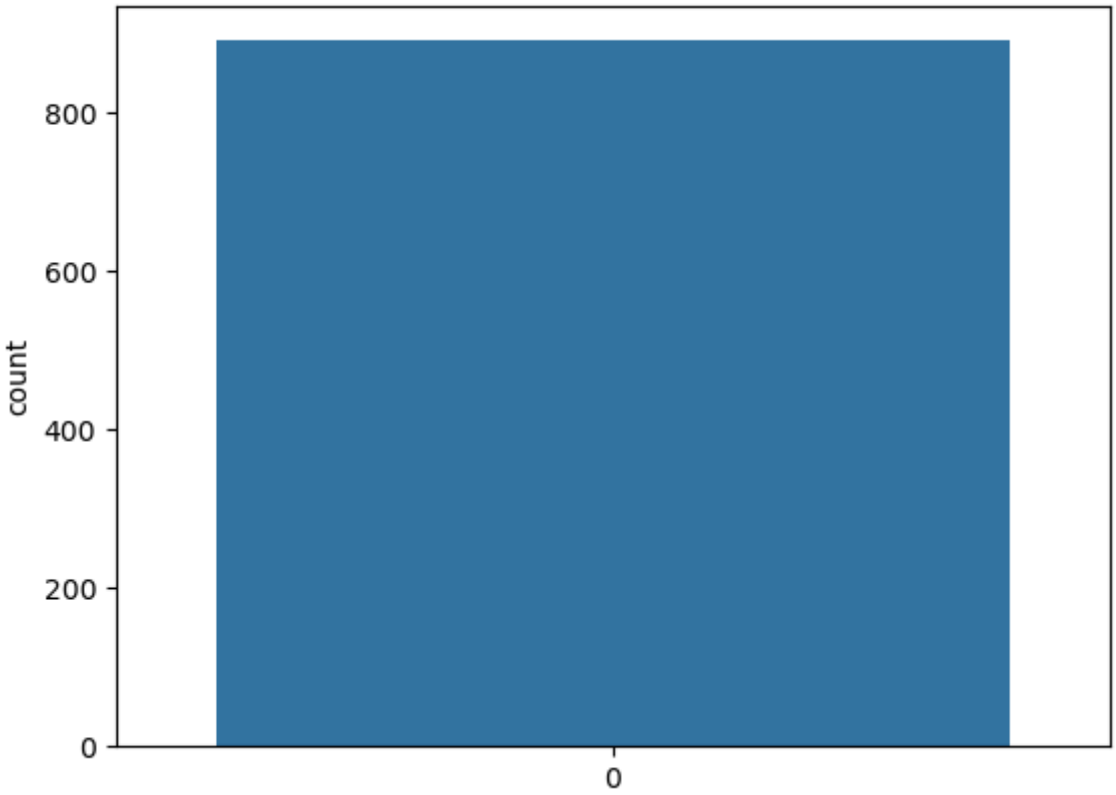
In [7]:
```
sns.countplot(data['Survived'])
```

Out[7]:
```
<Axes: ylabel='count'>
```



In [8]:
```
sns.countplot(data['Pclass'])
```

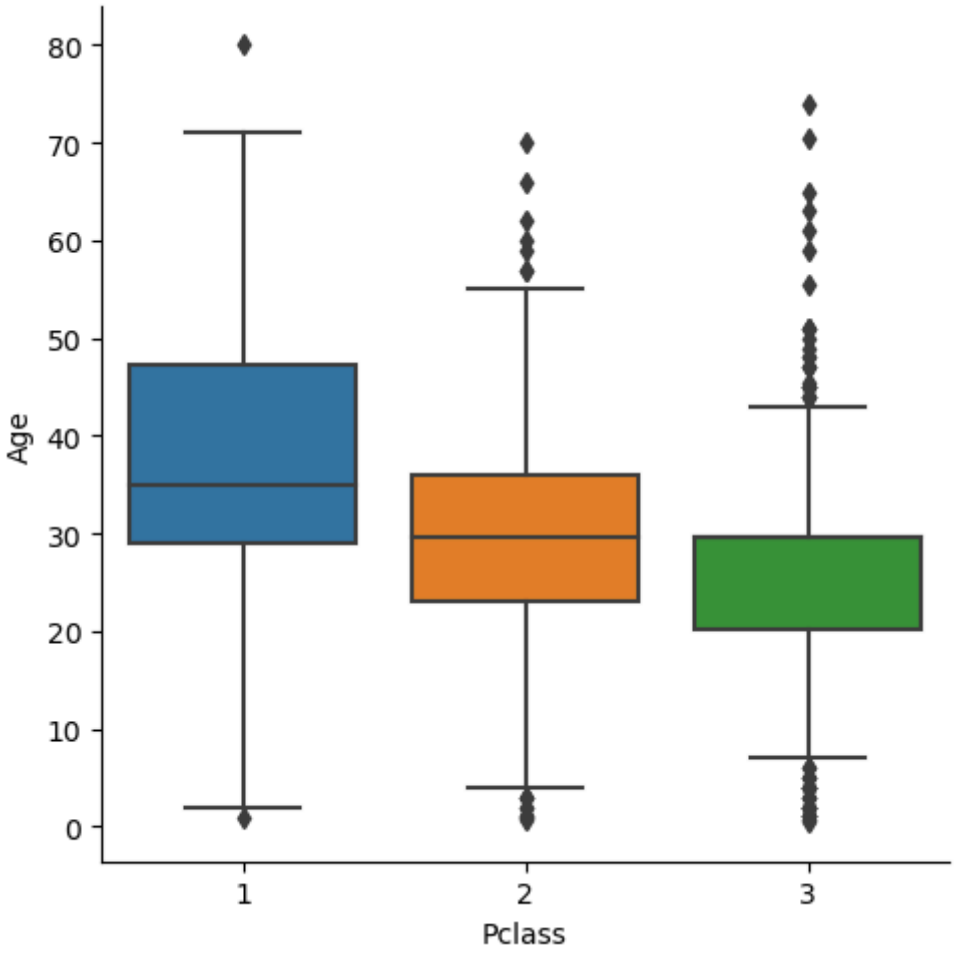
Out[8]:
```
<Axes: ylabel='count'>
```



In [9]:
```
sns.catplot(x= 'Pclass', y = 'Age', data=data, kind = 'box')
```

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

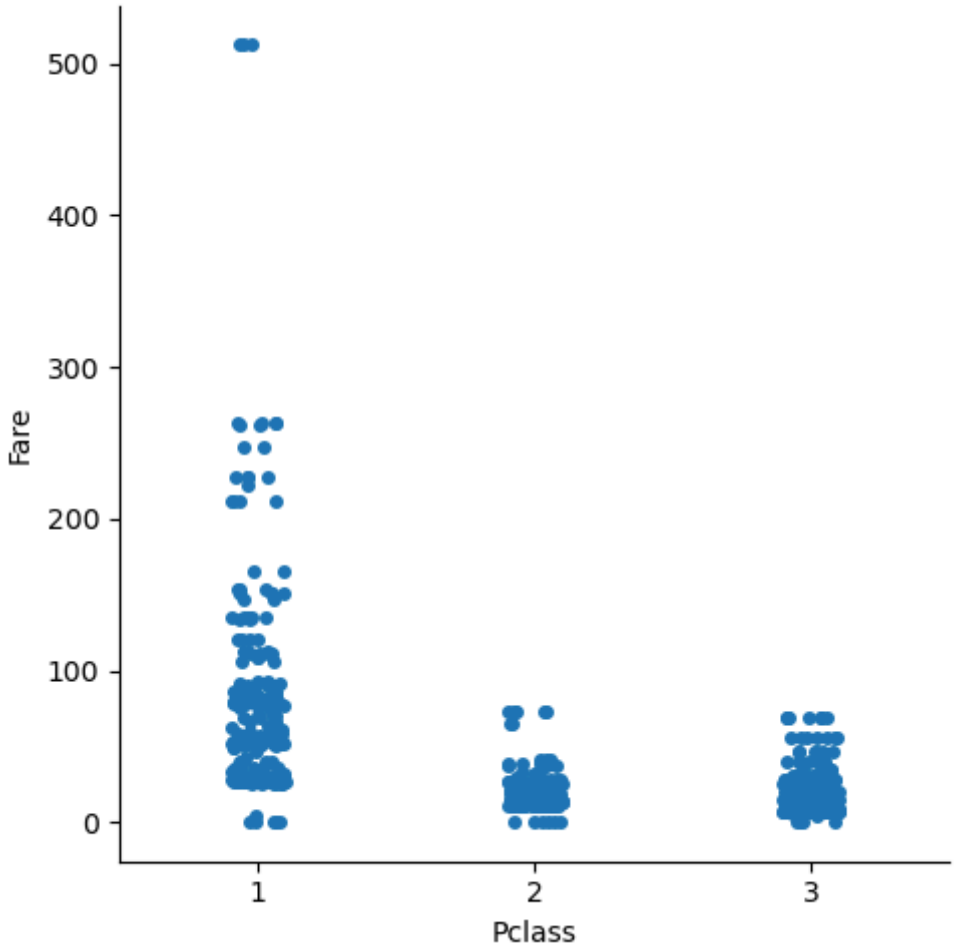
Out[9]:
```
<seaborn.axisgrid.FacetGrid at 0x13a2ac92f90>
```



In [10]:
```
sns.catplot(x= 'Pclass', y = 'Fare', data=data, kind = 'strip')
```

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

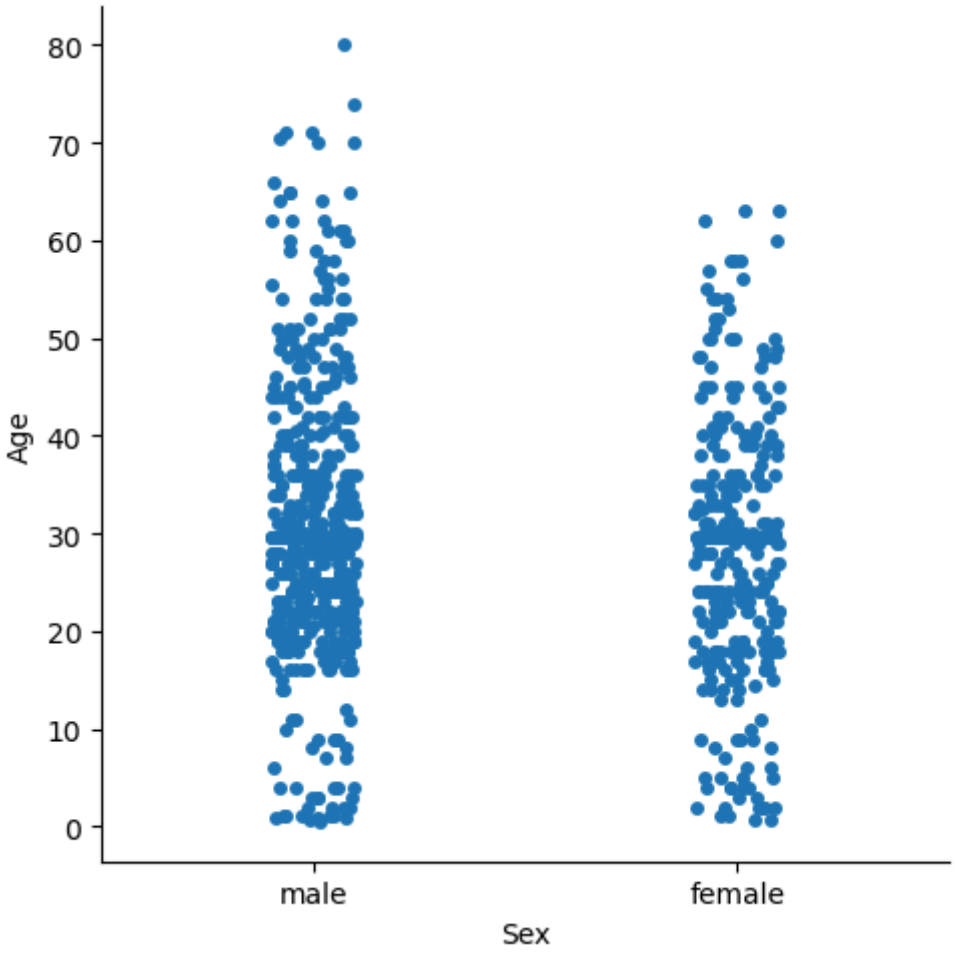
Out[10]:
```
<seaborn.axisgrid.FacetGrid at 0x13a2ad56210>
```



In [11]:
```
sns.catplot(x= 'Sex', y = 'Age', data=data, kind = 'strip')
```

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

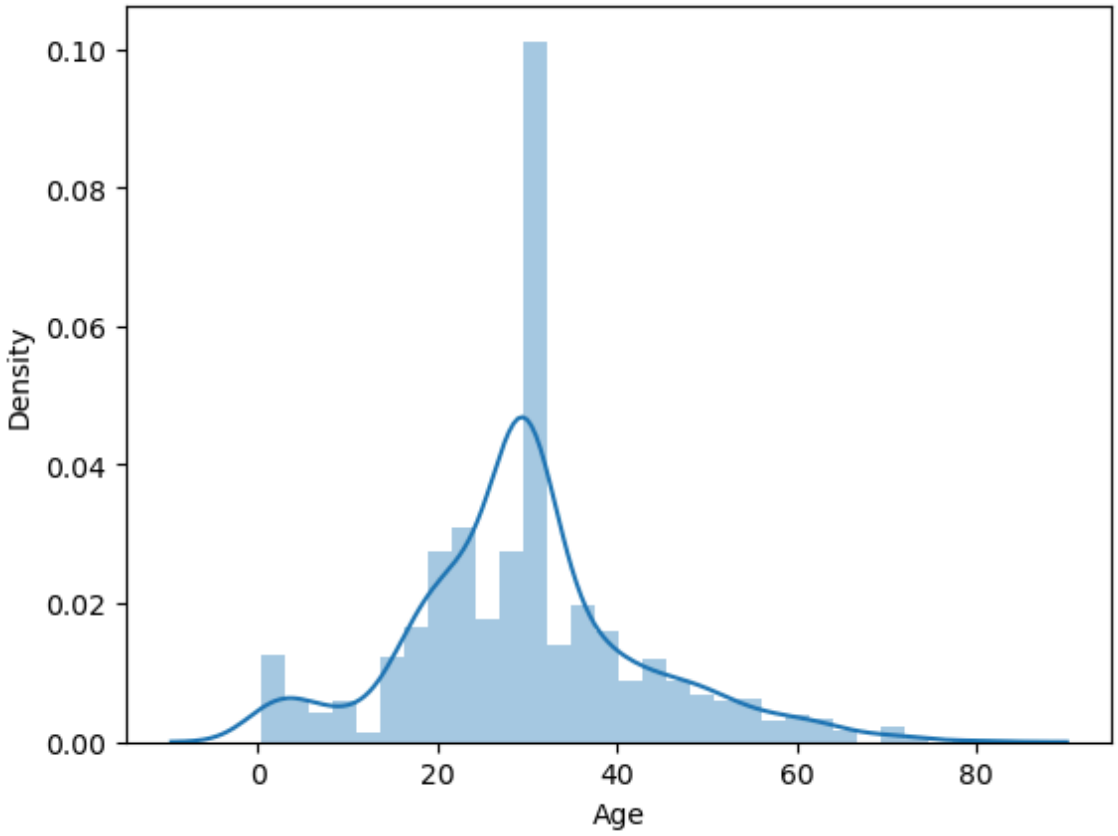
Out[11]:
```
<seaborn.axisgrid.FacetGrid at 0x13a2ad545d0>
```



In [12]:
```
sns.distplot(data['Age'])
```

```
C:\Users\Aniket\AppData\Local\Temp\ipykernel_19912\2317092479.py:1: UserWa
rning:

`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 wi
th
similar flexibility) or `histplot` (an axes-level function for histogram
s).

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'])
```

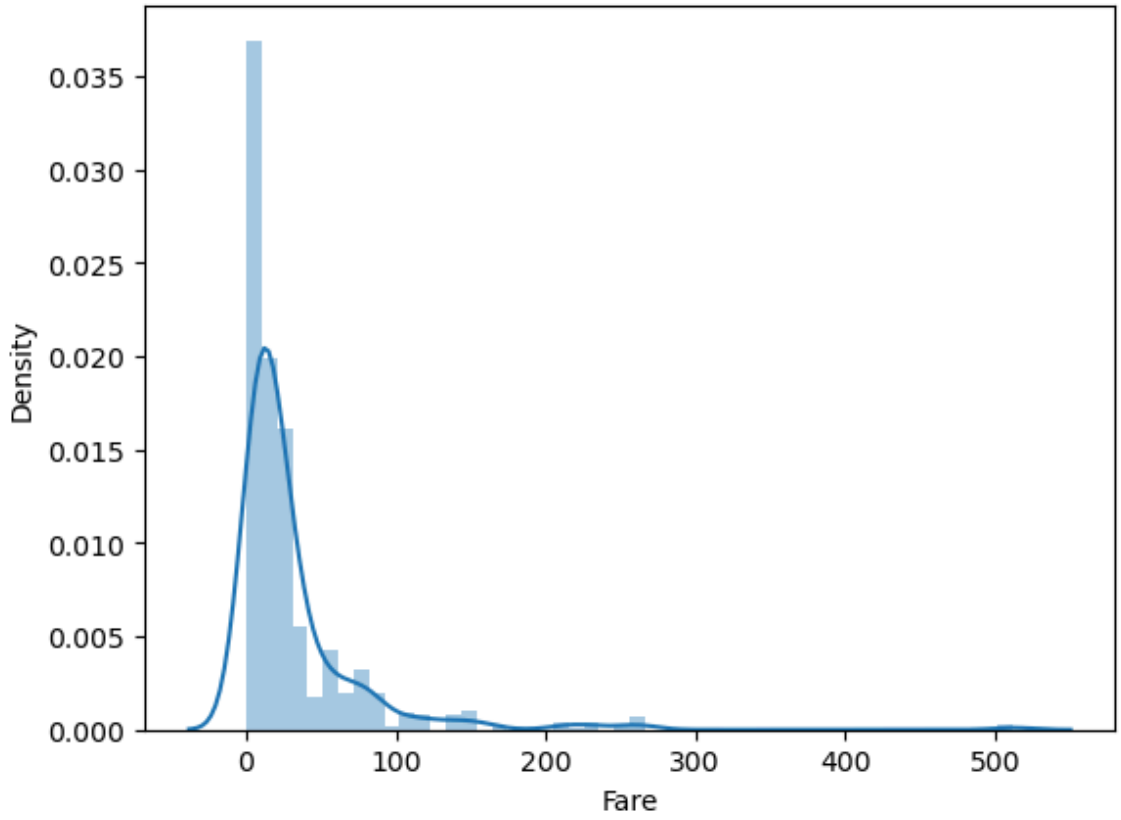
Out[12]:
```
<Axes: xlabel='Age', ylabel='Density'>
```



In [13]:
```
sns.distplot(data['Fare'])
```

```
C:\Users\Aniket\AppData\Local\Temp\ipykernel_19912\2921470011.py:1: UserWa
rning:

`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 wi
th
similar flexibility) or `histplot` (an axes-level function for histogram
s).

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['Fare'])
```

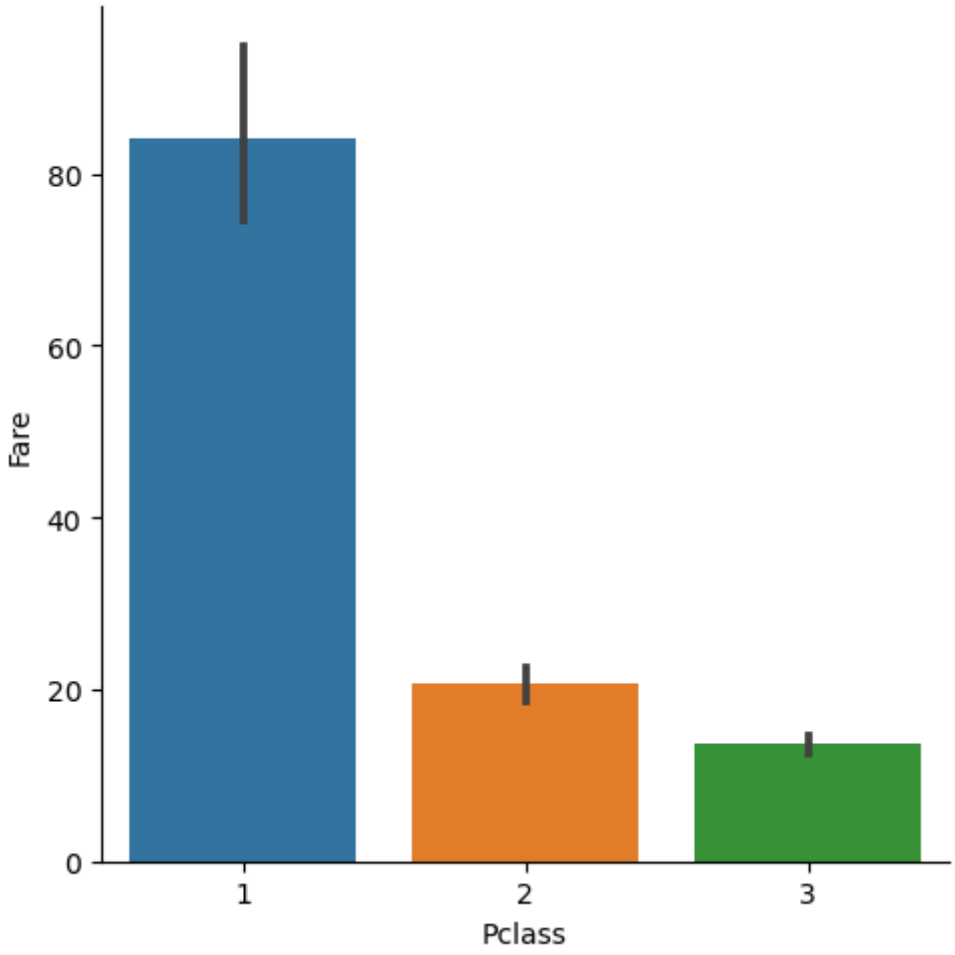
Out[13]:
```
<Axes: xlabel='Fare', ylabel='Density'>
```



In [14]:
```
sns.catplot(x='Pclass', y='Fare', data=data, kind='bar')
```

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

Out[14]:
```
<seaborn.axisgrid.FacetGrid at 0x13a2b052150>
```



In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: