Welcome to 30 Day ML | Day 7

Import Library

In [1]:
```
import pandas as pd
import seaborn as sns
```

Import Datasets 1: Titanic

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

In [4]:
```
titanic.head(3)
```

Out[4]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Emb
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	

Use Datasets 2: Hotel Bill & Tips

In [18]:
```
bill = sns.load_dataset('tips')
```

In [19]:
```
bill
```

Out[19]:

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4
...	...	...	...	...	...	...	...
239	29.03	5.92	Male	No	Sat	Dinner	3
240	27.18	2.00	Female	Yes	Sat	Dinner	2
241	22.67	2.00	Male	Yes	Sat	Dinner	2
242	17.82	1.75	Male	No	Sat	Dinner	2
243	18.78	3.00	Female	No	Thur	Dinner	2

244 rows × 7 columns

Use Datasets 3: USA Flights

In [8]:
```
flights = sns.load_dataset('flights')
```

In [10]:
```
flights
```

Out[10]:

	year	month	passengers
0	1949	Jan	112
1	1949	Feb	118
2	1949	Mar	132
3	1949	Apr	129
4	1949	May	121
...	...	...	...
139	1960	Aug	606
140	1960	Sep	508
141	1960	Oct	461
142	1960	Nov	390
143	1960	Dec	432

144 rows × 3 columns

Use Datasets 4: Iris Flower

In [11]:
```
irisflwr = sns.load_dataset('iris')
```

In [12]:
```
irisflwr
```

Out[12]:

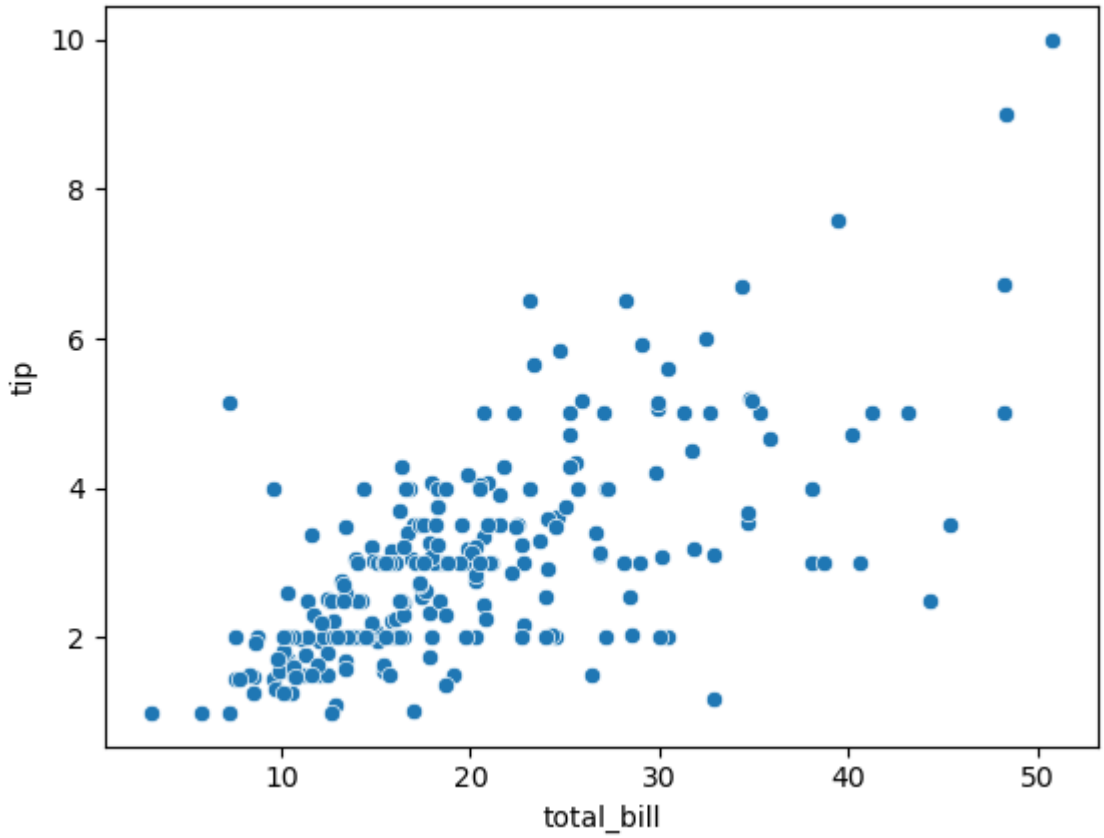
	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa
...	...	...	...	...	...
145	6.7	3.0	5.2	2.3	virginica
146	6.3	2.5	5.0	1.9	virginica
147	6.5	3.0	5.2	2.0	virginica
148	6.2	3.4	5.4	2.3	virginica
149	5.9	3.0	5.1	1.8	virginica

150 rows × 5 columns

1<Numerical – Numerical> | Scatterplot

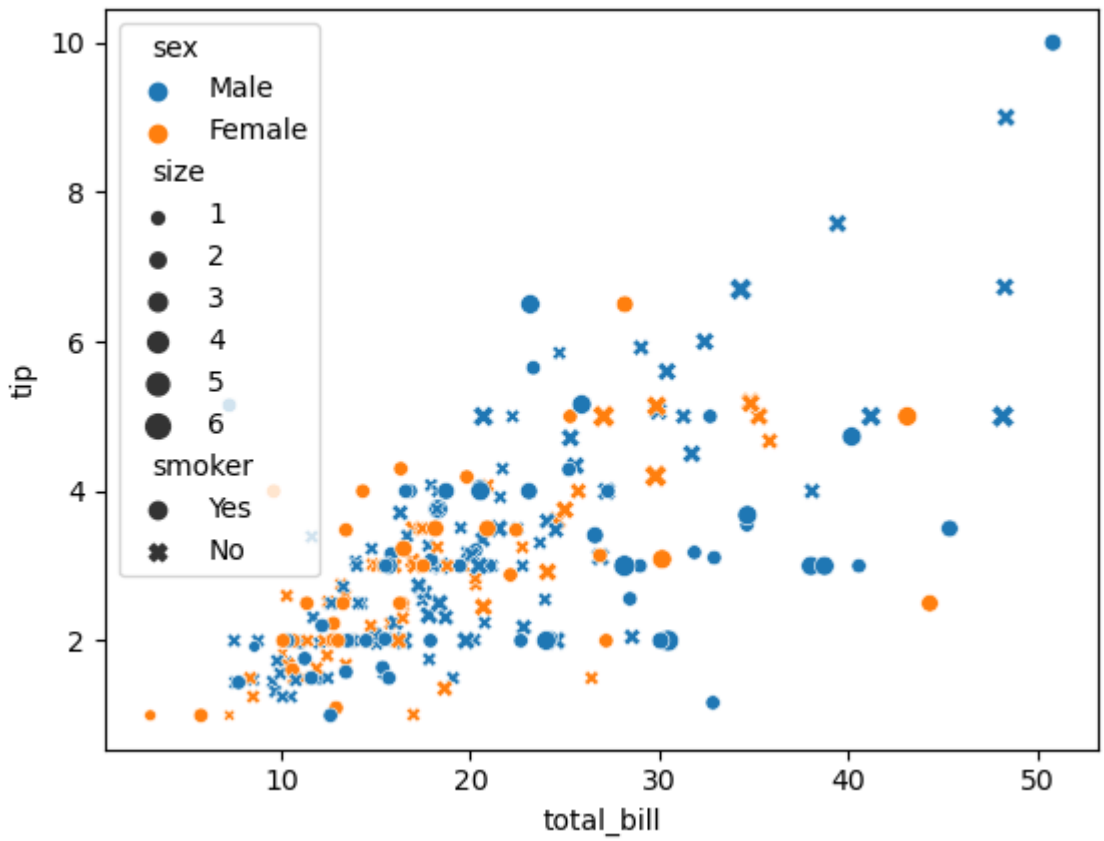
In [20]:
```
sns.scatterplot(x=bill['total_bill'],y=bill['tip'])
```

Out[20]: <Axes: xlabel='total_bill', ylabel='tip'>



In [26]:
```
sns.scatterplot(x=bill['total_bill'],y=bill['tip'],hue=bill['sex'],style=bill['smoker
```

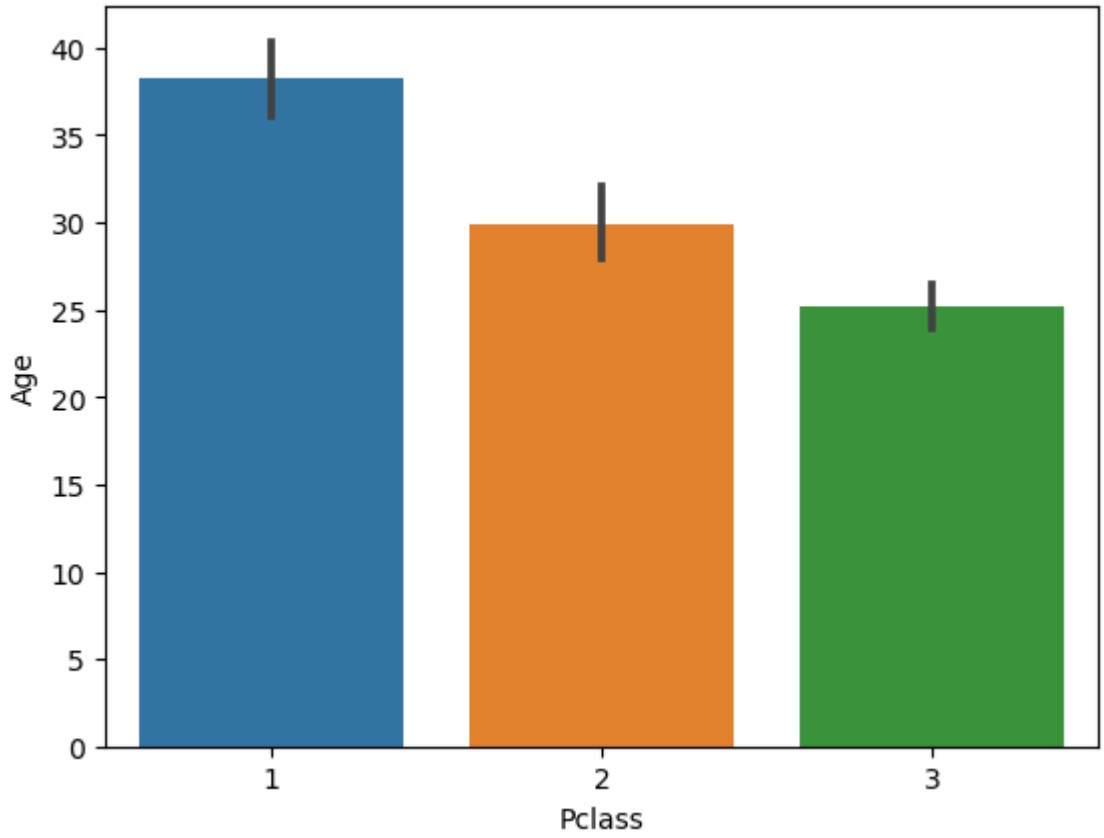
Out[26]: <Axes: xlabel='total_bill', ylabel='tip'>



2<Numerical –Categorical> | Bar plot

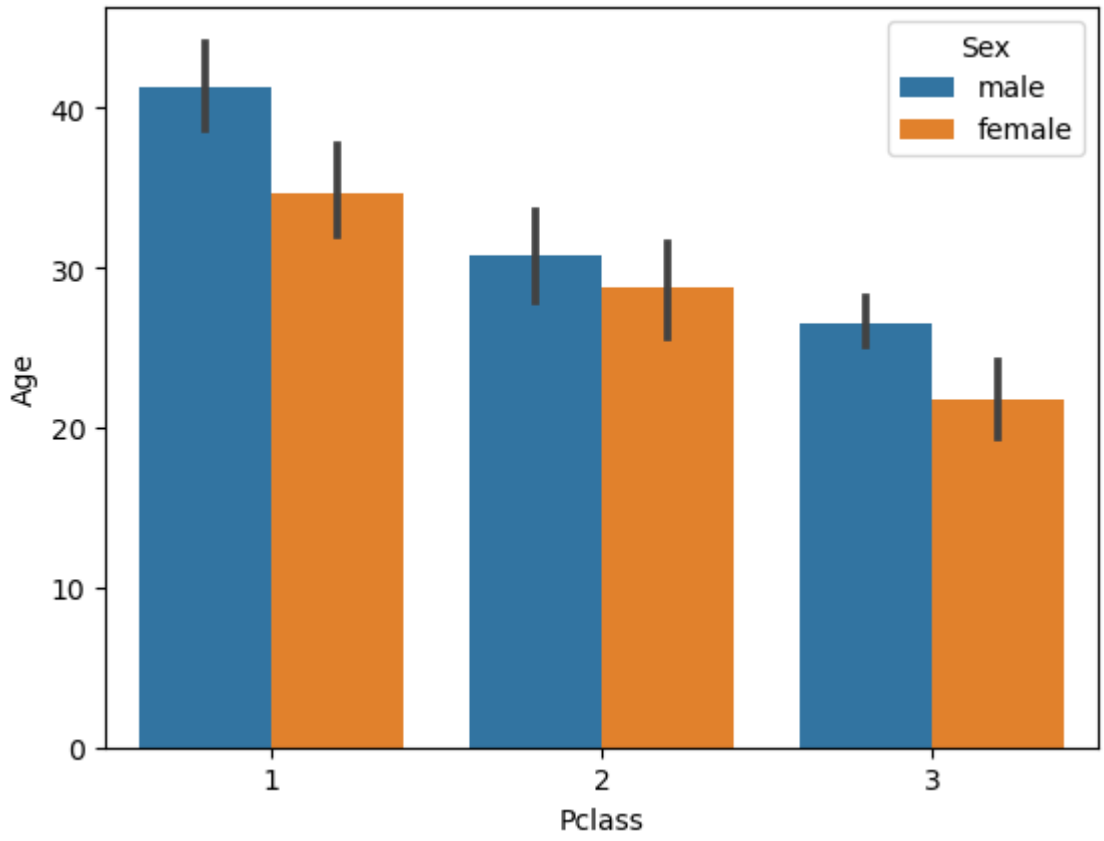
In [28]:
```
sns.barplot(x=titanic['Pclass'],y=titanic['Age'])
```

Out[28]: <Axes: xlabel='Pclass', ylabel='Age'>



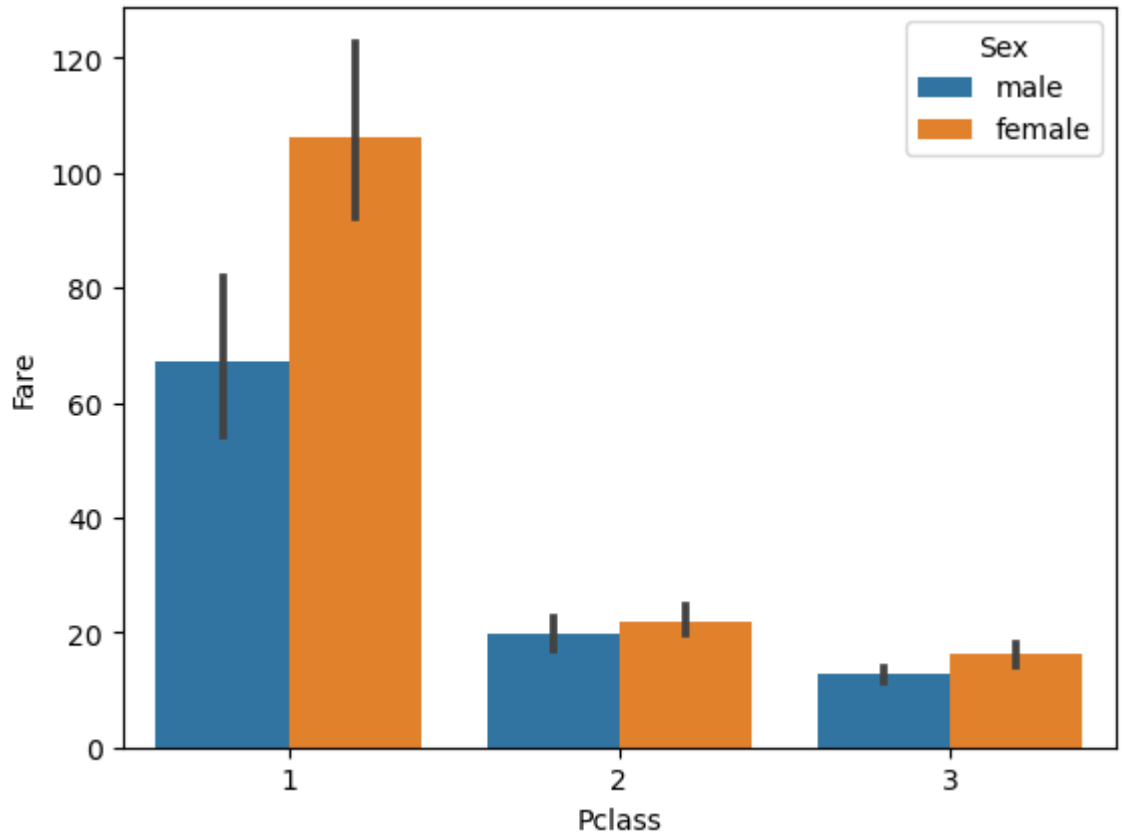
In [29]:
```
sns.barplot(x=titanic['Pclass'],y=titanic['Age'],hue=titanic['Sex'])
```

Out[29]: <Axes: xlabel='Pclass', ylabel='Age'>



In [30]:
```
sns.barplot(x=titanic['Pclass'],y=titanic['Fare'],hue=titanic['Sex'])
```

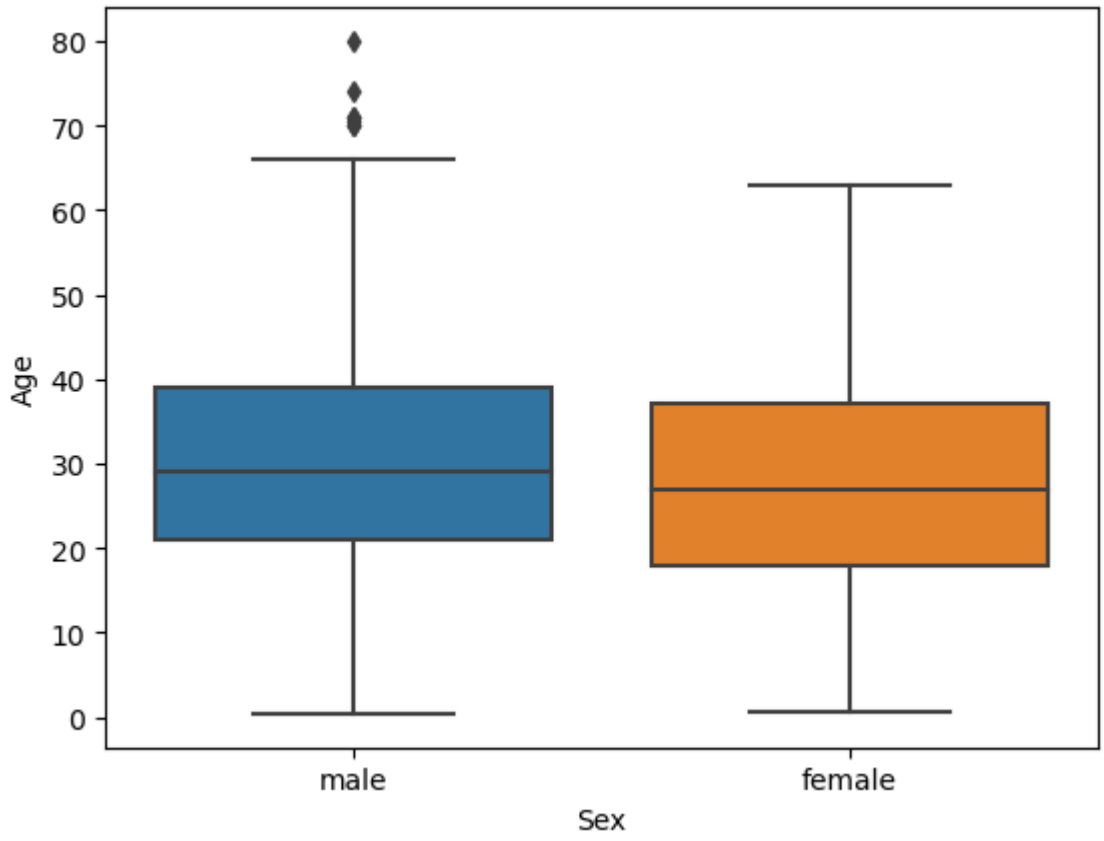
Out[30]: <Axes: xlabel='Pclass', ylabel='Fare'>



3<Numerical –Categorical> | Box Plot

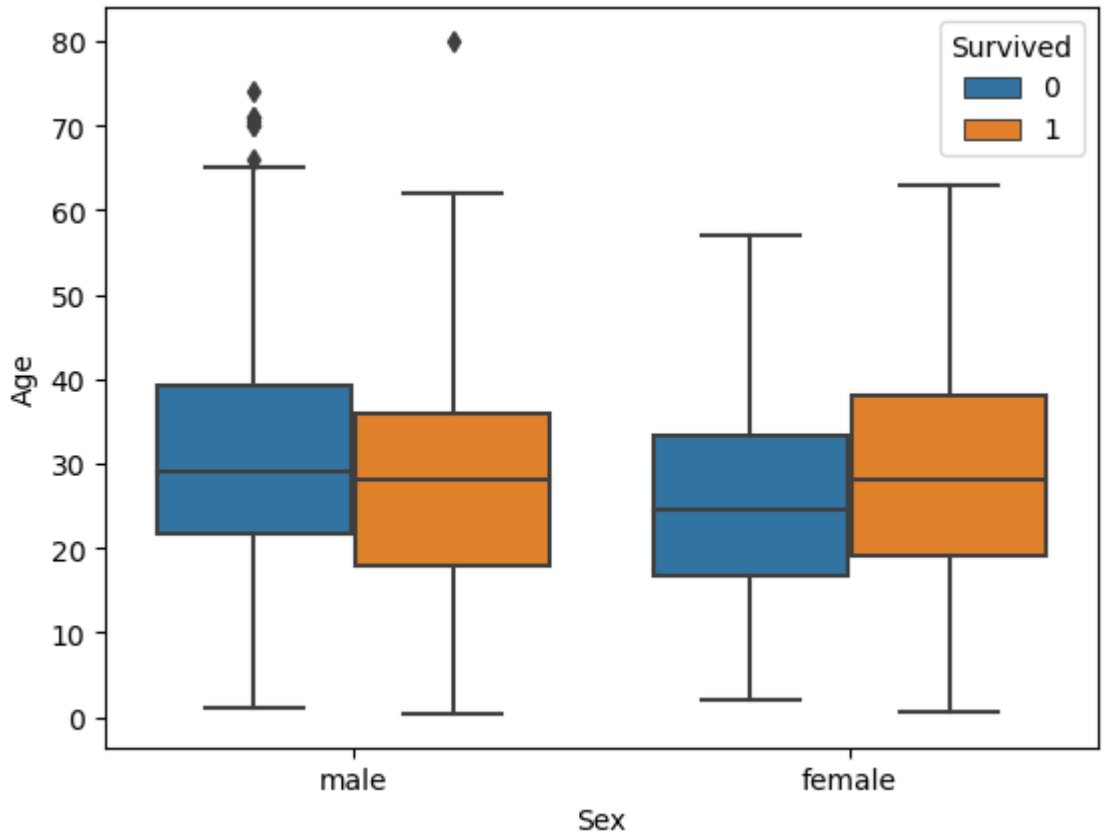
In [31]:
```
sns.boxplot(x=titanic['Sex'],y=titanic['Age'])
```

Out[31]: <Axes: xlabel='Sex', ylabel='Age'>



In [32]:
```
sns.boxplot(x=titanic['Sex'],y=titanic['Age'],hue=titanic['Survived'])
```

Out[32]:
```
<Axes: xlabel='Sex', ylabel='Age'>
```

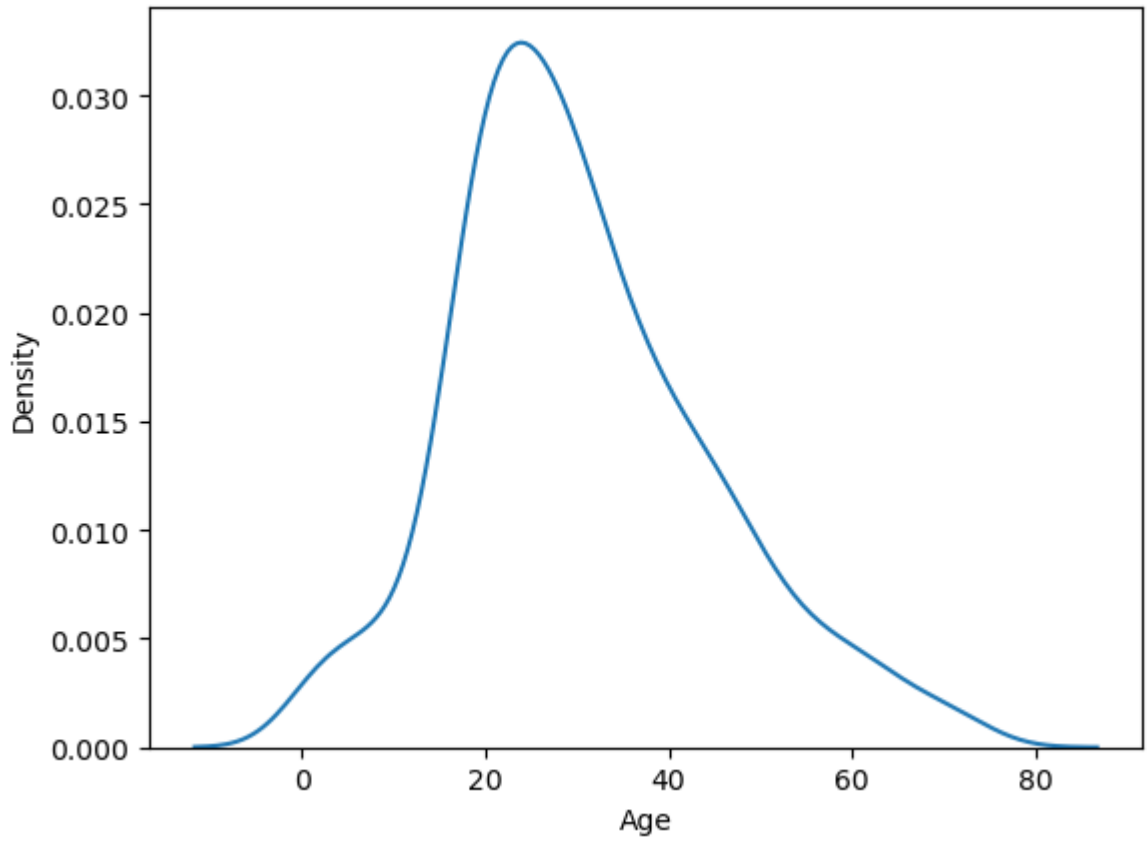


4<Numerical –Categorical> | Dist Plot

In [34]:
```
sns.distplot(titanic[titanic['Survived']==0]['Age'],hist=False)
```

```
C:\Users\ASUS\AppData\Local\Temp\ipykernel_4560\2677591329.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 `kdeplot` (an axes-level function for kernel density plots).

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(titanic[titanic['Survived']==0]['Age'],hist=False)
```

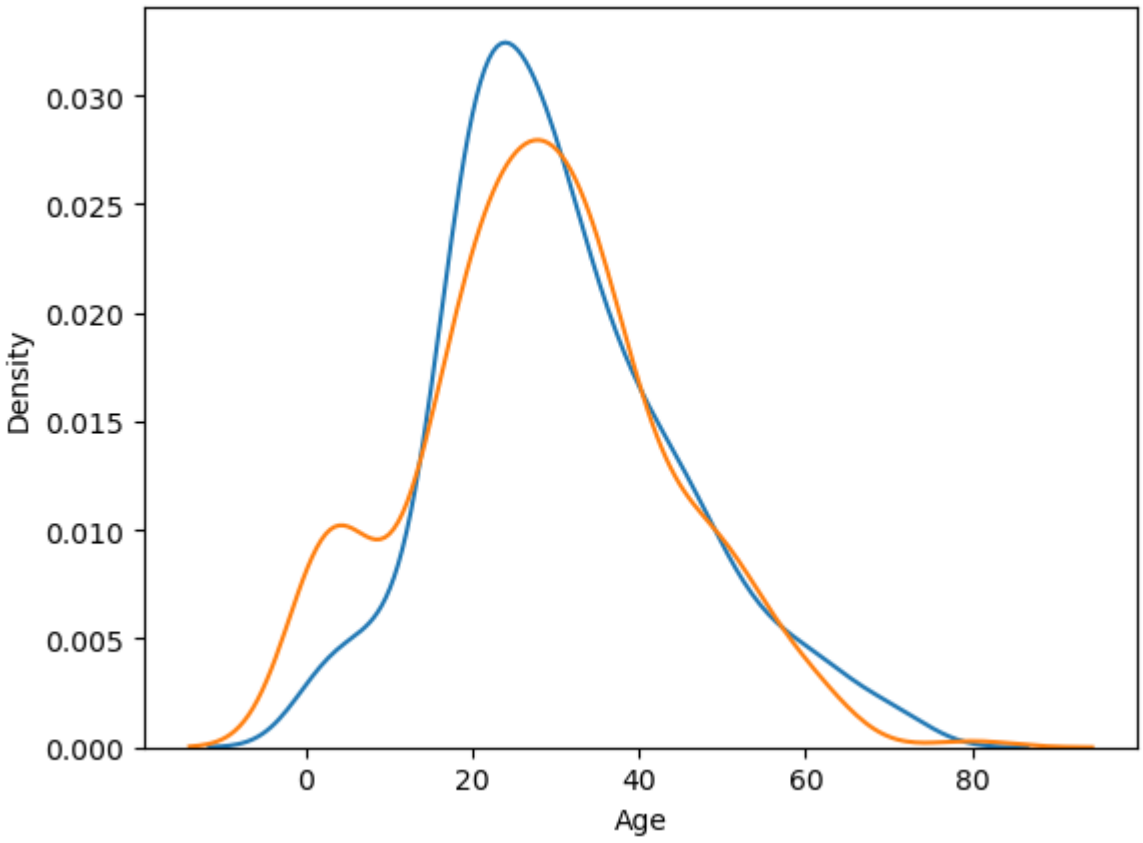
Out[34]:
```
<Axes: xlabel='Age', ylabel='Density'>
```



In [35]:
```
sns.distplot(titanic[titanic['Survived']==0]['Age'],hist=False)
sns.distplot(titanic[titanic['Survived']==1]['Age'],hist=False)
```

```
C:\Users\ASUS\AppData\Local\Temp\ipykernel_4560\1261424998.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 `kdeplot` (an axes-level function for kernel density plots).

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(titanic[titanic['Survived']==0]['Age'],hist=False)
C:\Users\ASUS\AppData\Local\Temp\ipykernel_4560\1261424998.py:2: 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 `kdeplot` (an axes-level function for kernel density plots).

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(titanic[titanic['Survived']==1]['Age'],hist=False)
```

Out[35]:
```
<Axes: xlabel='Age', ylabel='Density'>
```



5<Categorical – Categorical> | HeatMap

In [45]:
```
titanic
```

Out[45]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Er
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	
...	...	...	...	...	...	...	...	...	...	...	...	...
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	NaN	
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	B42	
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	NaN	
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C148	
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	NaN	

891 rows × 12 columns

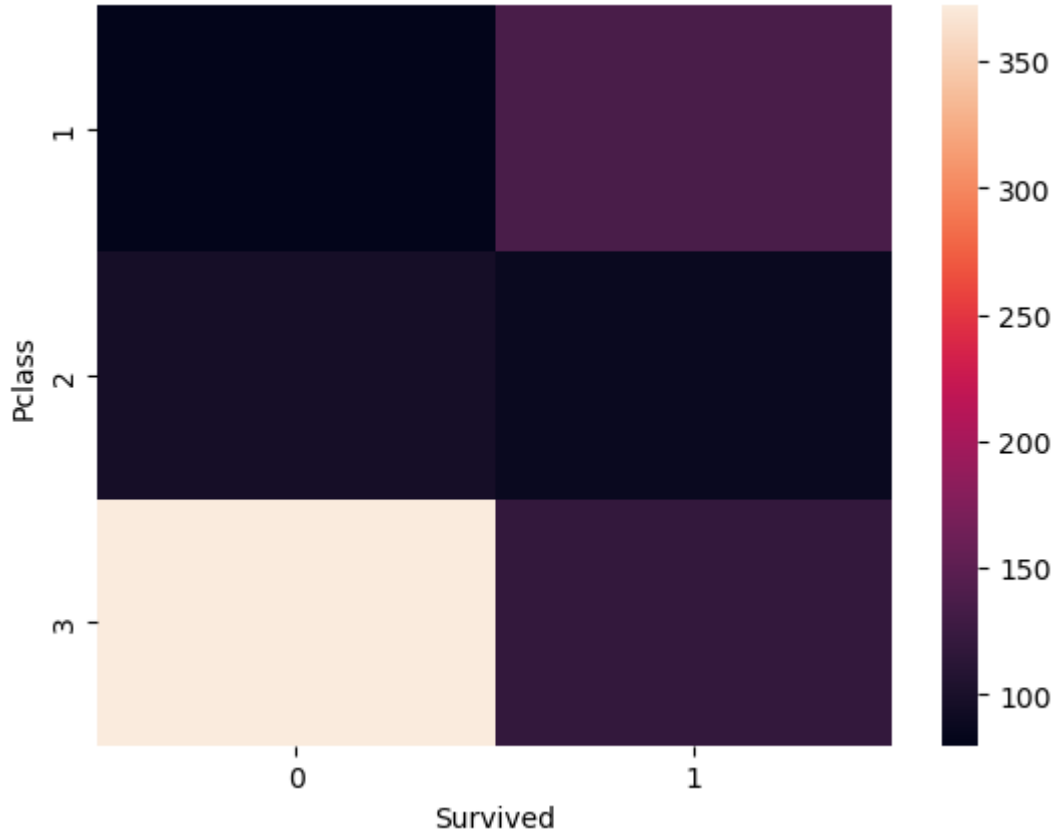
In [37]:
```
pd.crosstab(titanic['Pclass'],titanic['Survived'])
```

Out[37]:

Survived	0	1
Pclass		
1	80	136
2	97	87
3	372	119

In [38]:
```
sns.heatmap(pd.crosstab(titanic['Pclass'],titanic
['Survived']))
```

Out[38]: <Axes: xlabel='Survived', ylabel='Pclass'>



6<Categorical – Categorical> | ClusterMap

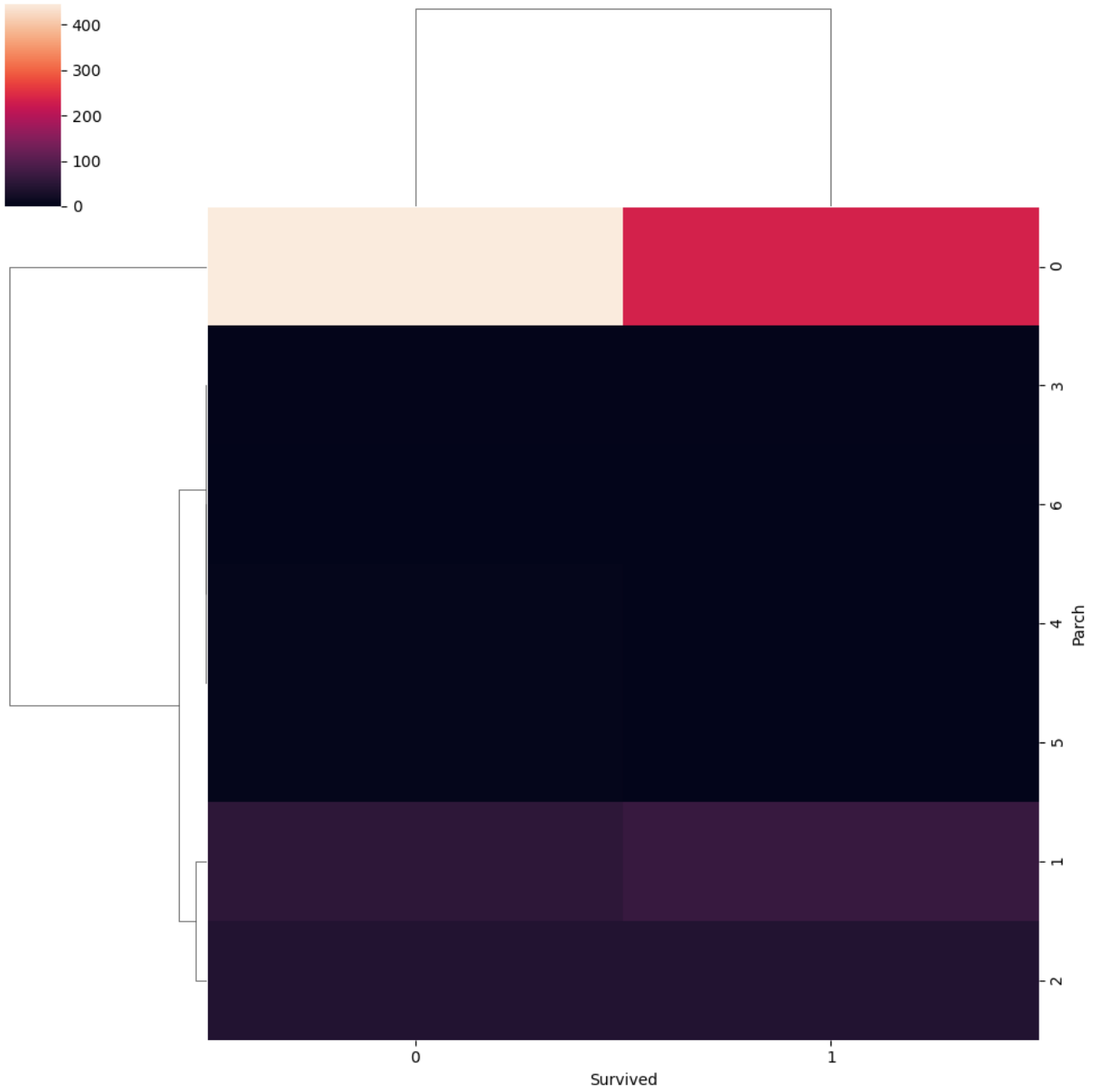
In [46]:
```
pd.crosstab(titanic['SibSp'],titanic['Survived'])
```

Out[46]:

Survived	0	1
SibSp		
0	398	210
1	97	112
2	15	13
3	12	4
4	15	3
5	5	0
8	7	0

In [43]:
```
sns.clustermap(pd.crosstab(titanic['Parch'],titanic['Survived']))
```

Out[43]:
```
<seaborn.matrix.ClusterGrid at 0x2bc24596710>
```



7<Numerical – Numerical -Categorical> | Pare Plot

In [44]:
```
irisflwr.head()
```

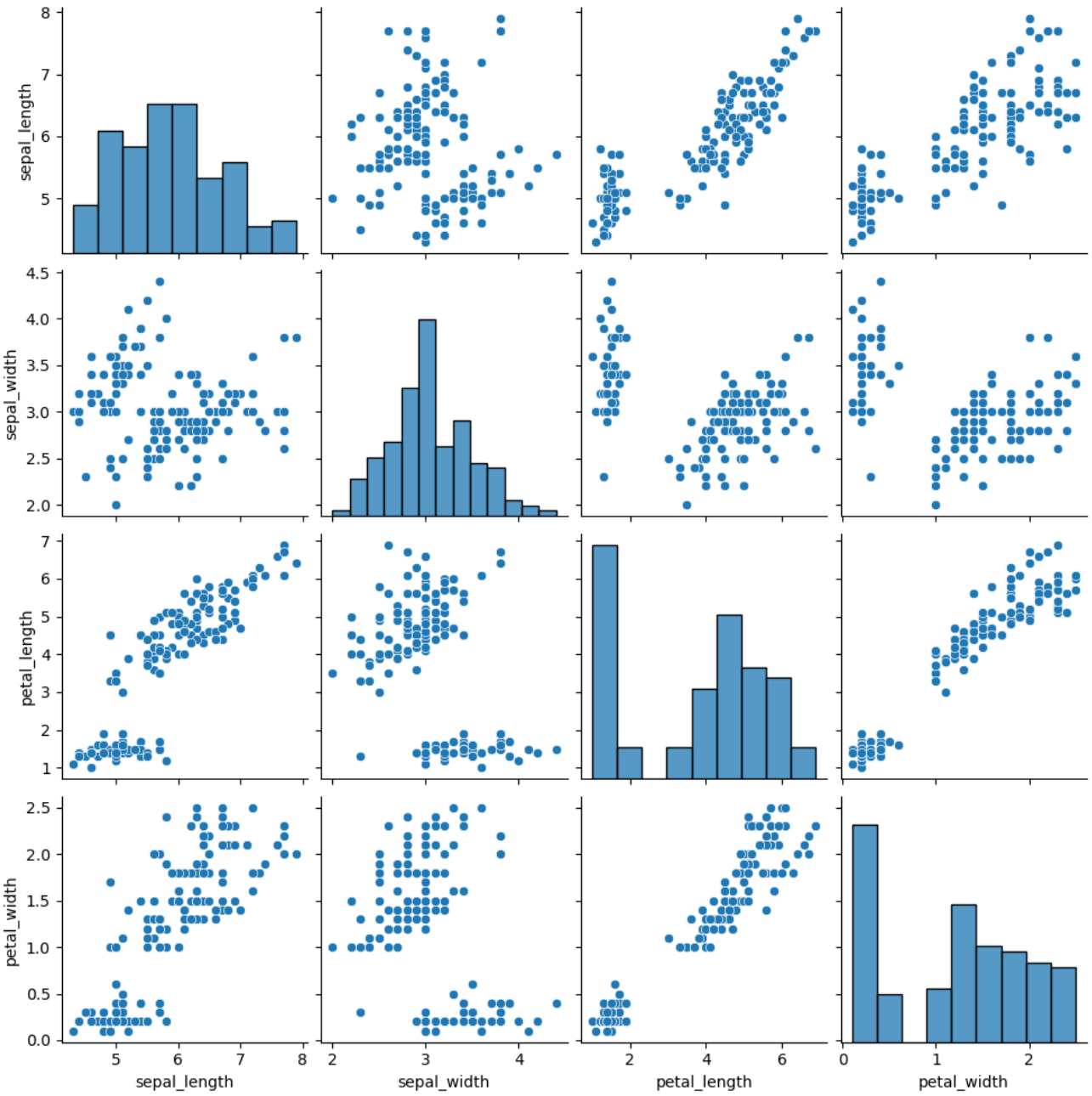
Out[44]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

In [47]:
```
sns.pairplot(irisflwr)
```

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

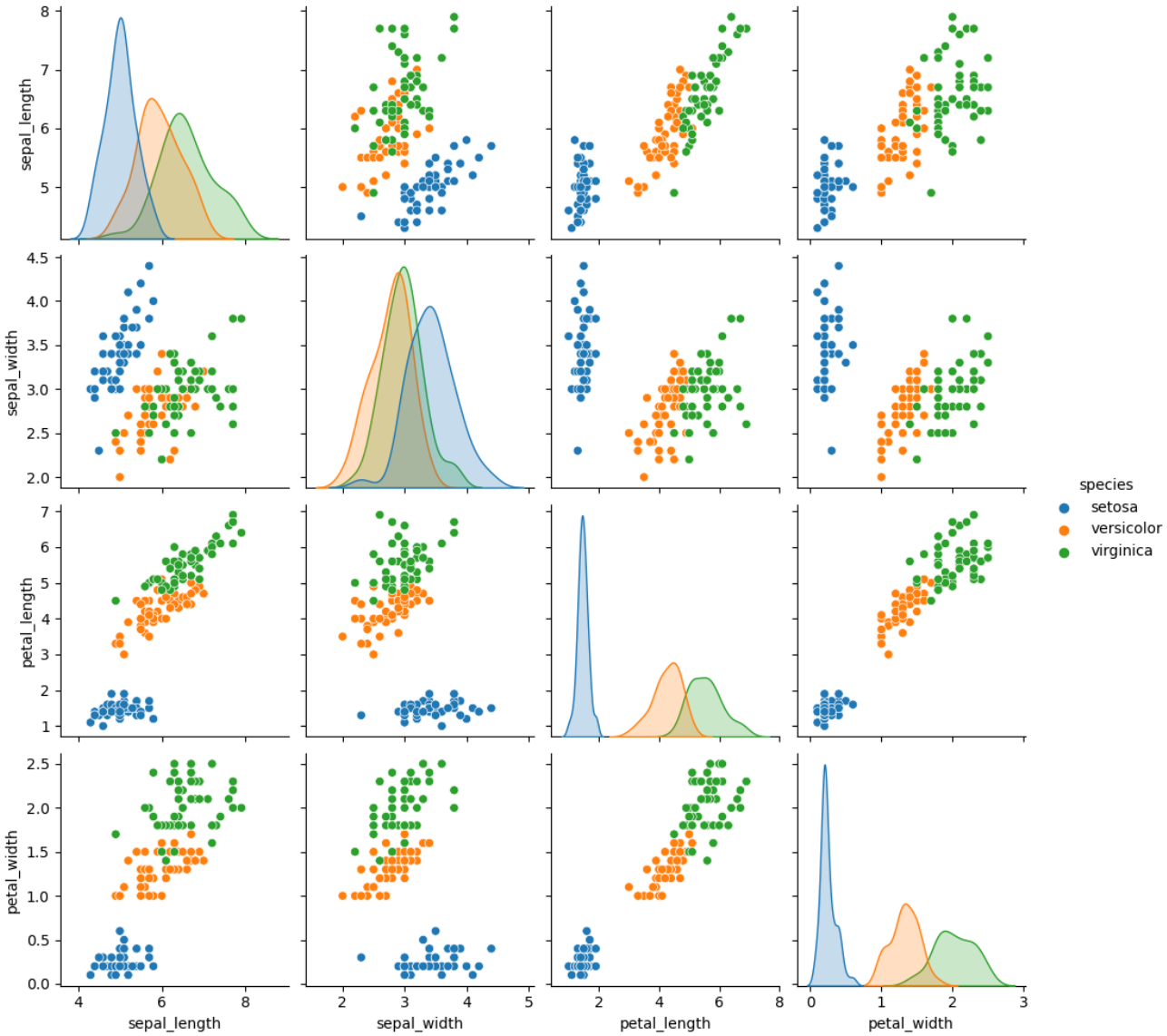
Out[47]:
```
<seaborn.axisgrid.PairGrid at 0x2bc2464bad0>
```



In [48]:
```
sns.pairplot(irisflwr,hue='species')
```

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

Out[48]:
```
<seaborn.axisgrid.PairGrid at 0x2bc2592f850>
```



In [ ]: