In [1]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

In [2]:
```
a=pd.read_csv('pen.csv')
a
```

Out[2]:

	island	bill_length_mm	bill_depth_mm	flipper_length_mm	body_mass_g
0	Torgersen	39.1	18.7	181.0	3750.0
1	Torgersen	39.5	17.4	186.0	3800.0
2	Torgersen	40.3	18.0	195.0	3250.0
3	Torgersen	NaN	NaN	NaN	NaN
4	Torgersen	36.7	19.3	193.0	3450.0
...	...	...	...	...	...
339	Dream	55.8	19.8	207.0	4000.0
340	Dream	43.5	18.1	202.0	3400.0
341	Dream	49.6	18.2	193.0	3775.0
342	Dream	50.8	19.0	210.0	4100.0
343	Dream	50.2	18.7	198.0	3775.0

344 rows × 7 columns

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

Out[3]:

	island	bill_length_mm	bill_depth_mm	flipper_length_mm	body_mass_g	G
0	Torgersen	39.1	18.7	181.0	3750.0	
1	Torgersen	39.5	17.4	186.0	3800.0	
2	Torgersen	40.3	18.0	195.0	3250.0	
3	Torgersen	NaN	NaN	NaN	NaN	
4	Torgersen	36.7	19.3	193.0	3450.0	

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

Out[4]:

	island	bill_length_mm	bill_depth_mm	flipper_length_mm	body_mass_g	G
339	Dream	55.8	19.8	207.0	4000.0	
340	Dream	43.5	18.1	202.0	3400.0	f
341	Dream	49.6	18.2	193.0	3775.0	
342	Dream	50.8	19.0	210.0	4100.0	
343	Dream	50.2	18.7	198.0	3775.0	f

In [5]:
```
a.shape
```

Out[5]:
```
(344, 7)
```

In [6]:
```
a.size
```

Out[6]:
```
2408
```

In [7]:
```
a.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 344 entries, 0 to 343
Data columns (total 7 columns):
 #   Column             Non-Null Count  Dtype  
---  ------             --------------  -----  
 0   island             344 non-null    object 
 1   bill_length_mm     342 non-null    float64
 2   bill_depth_mm      342 non-null    float64
 3   flipper_length_mm  342 non-null    float64
 4   body_mass_g        342 non-null    float64
 5   Gender             333 non-null    object 
 6   species            344 non-null    object 
dtypes: float64(4), object(3)
memory usage: 18.9+ KB
```

In [8]:
```
a.describe()
```

Out[8]:

	bill_length_mm	bill_depth_mm	flipper_length_mm	body_mass_g
count	342.000000	342.000000	342.000000	342.000000
mean	43.921930	17.151170	200.915205	4201.754386
std	5.459584	1.974793	14.061714	801.954536
min	32.100000	13.100000	172.000000	2700.000000
25%	39.225000	15.600000	190.000000	3550.000000
50%	44.450000	17.300000	197.000000	4050.000000
75%	48.500000	18.700000	213.000000	4750.000000
max	59.600000	21.500000	231.000000	6300.000000

In [9]:
```
a_adelie=a[a['species']=='Adelie']
a_adelie
```

Out[9]:

	island	bill_length_mm	bill_depth_mm	flipper_length_mm	body_mass_g
0	Torgersen	39.1	18.7	181.0	3750.0
1	Torgersen	39.5	17.4	186.0	3800.0
2	Torgersen	40.3	18.0	195.0	3250.0
3	Torgersen	NaN	NaN	NaN	NaN
4	Torgersen	36.7	19.3	193.0	3450.0
...	...	...	...	...	...
147	Dream	36.6	18.4	184.0	3475.0
148	Dream	36.0	17.8	195.0	3450.0
149	Dream	37.8	18.1	193.0	3750.0
150	Dream	36.0	17.1	187.0	3700.0
151	Dream	41.5	18.5	201.0	4000.0

152 rows × 7 columns

In [11]:
```
a_gentoo=a[a['species']=='Gentoo']
a_gentoo
```

Out[11]:

	island	bill_length_mm	bill_depth_mm	flipper_length_mm	body_mass_g	G
152	Biscoe	46.1	13.2	211.0	4500.0	f
153	Biscoe	50.0	16.3	230.0	5700.0	
154	Biscoe	48.7	14.1	210.0	4450.0	f
155	Biscoe	50.0	15.2	218.0	5700.0	
156	Biscoe	47.6	14.5	215.0	5400.0	
...	...	...	...	...	...	
271	Biscoe	NaN	NaN	NaN	NaN	
272	Biscoe	46.8	14.3	215.0	4850.0	f
273	Biscoe	50.4	15.7	222.0	5750.0	
274	Biscoe	45.2	14.8	212.0	5200.0	f
275	Biscoe	49.9	16.1	213.0	5400.0	

124 rows × 7 columns

In [12]:
```
a_chinstrap=a[a['species']=='Chinstrap']
```

```
a_chinstrap
```

Out[12]:

	island	bill_length_mm	bill_depth_mm	flipper_length_mm	body_mass_g	G
276	Dream	46.5	17.9	192.0	3500.0	f
277	Dream	50.0	19.5	196.0	3900.0	
278	Dream	51.3	19.2	193.0	3650.0	
279	Dream	45.4	18.7	188.0	3525.0	f
280	Dream	52.7	19.8	197.0	3725.0	
...	...	...	...	...	...	...
339	Dream	55.8	19.8	207.0	4000.0	
340	Dream	43.5	18.1	202.0	3400.0	f
341	Dream	49.6	18.2	193.0	3775.0	
342	Dream	50.8	19.0	210.0	4100.0	
343	Dream	50.2	18.7	198.0	3775.0	f

68 rows × 7 columns

In [13]:
```
a_adelie.shape
```

Out[13]:
```
(152, 7)
```

In [14]:
```
a_gentoo.shape
```

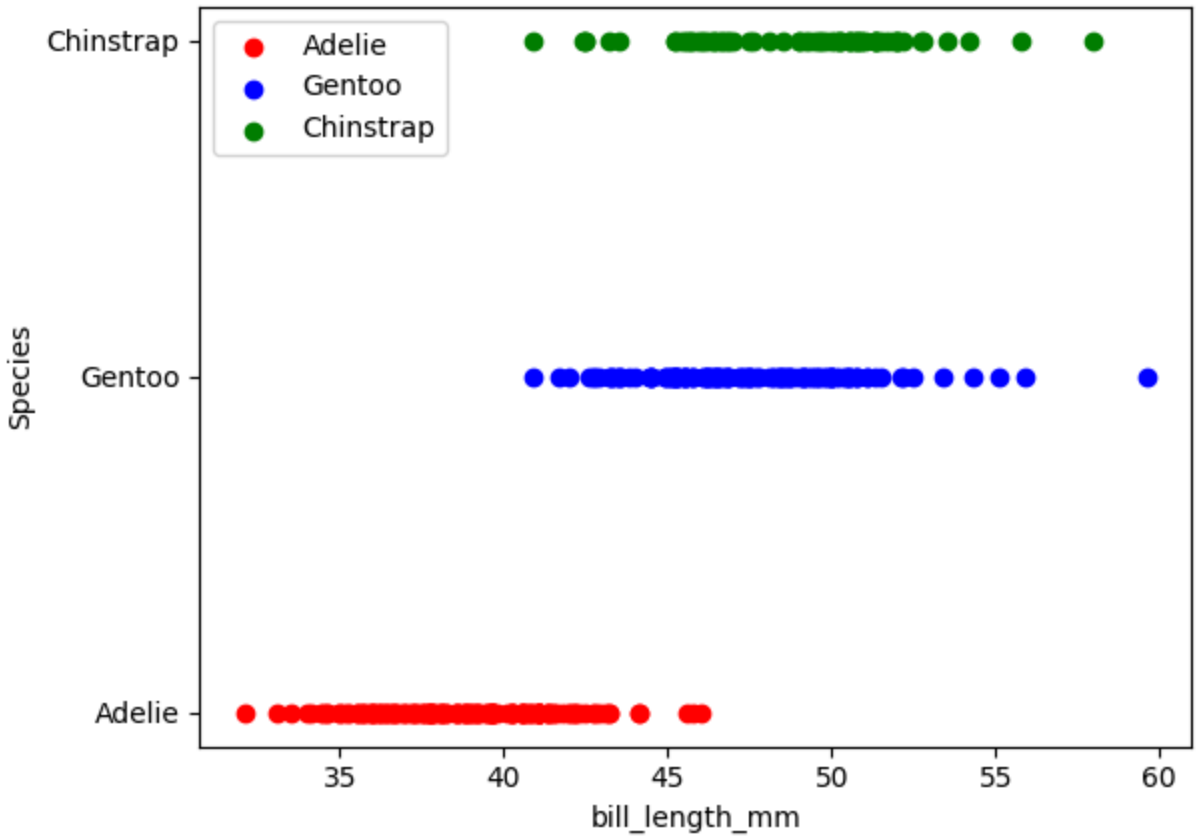
Out[14]:
```
(124, 7)
```

In [15]:
```
a_chinstrap.shape
```

Out[15]:
```
(68, 7)
```

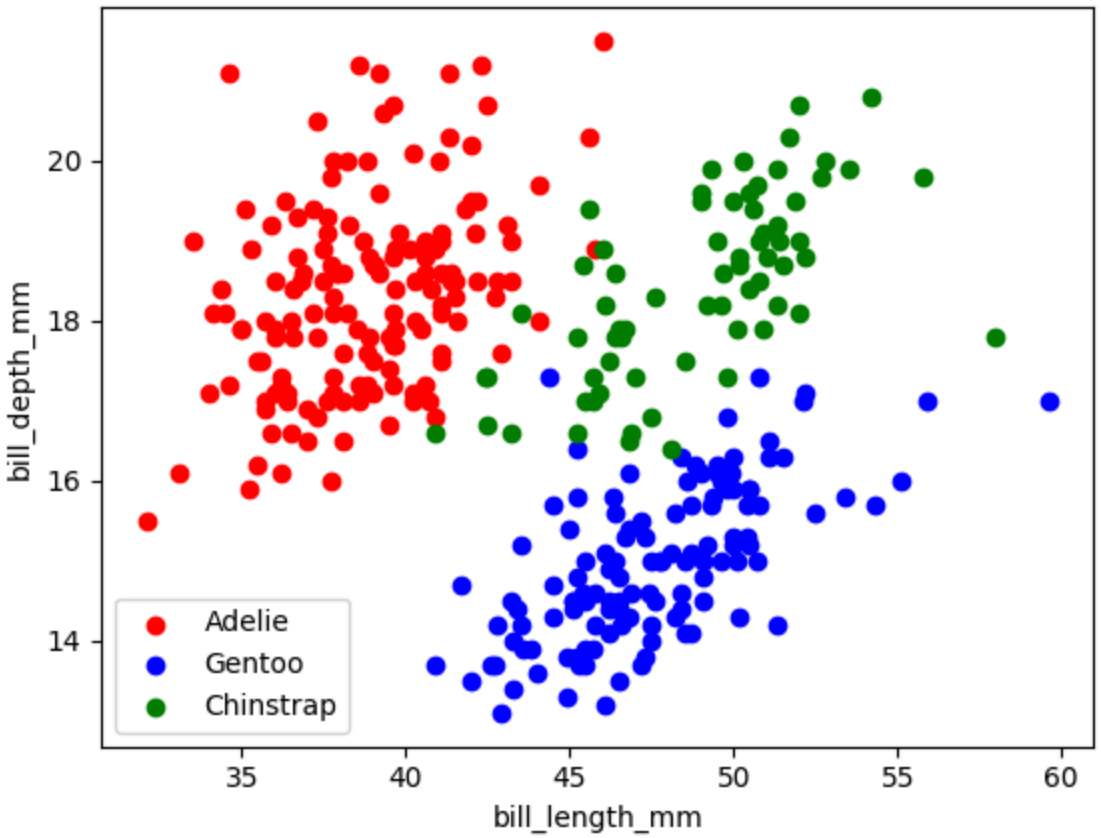
In [31]:
```
plt.scatter(a_adelie['bill_length_mm'],np.zeros_like(a_adelie['bill_length_mm'
plt.scatter(a_gentoo['bill_length_mm'],np.ones_like(a_gentoo['bill_length_mm']
plt.scatter(a_chinstrap['bill_length_mm'],np.full_like(a_chinstrap['bill_lengt
plt.xlabel('bill_length_mm')
plt.ylabel('Species')
plt.yticks([0,1,2],['Adelie','Gentoo','Chinstrap'])
plt.legend()
```

Out[31]:
```
<matplotlib.legend.Legend at 0x2e6147f20d0>
```



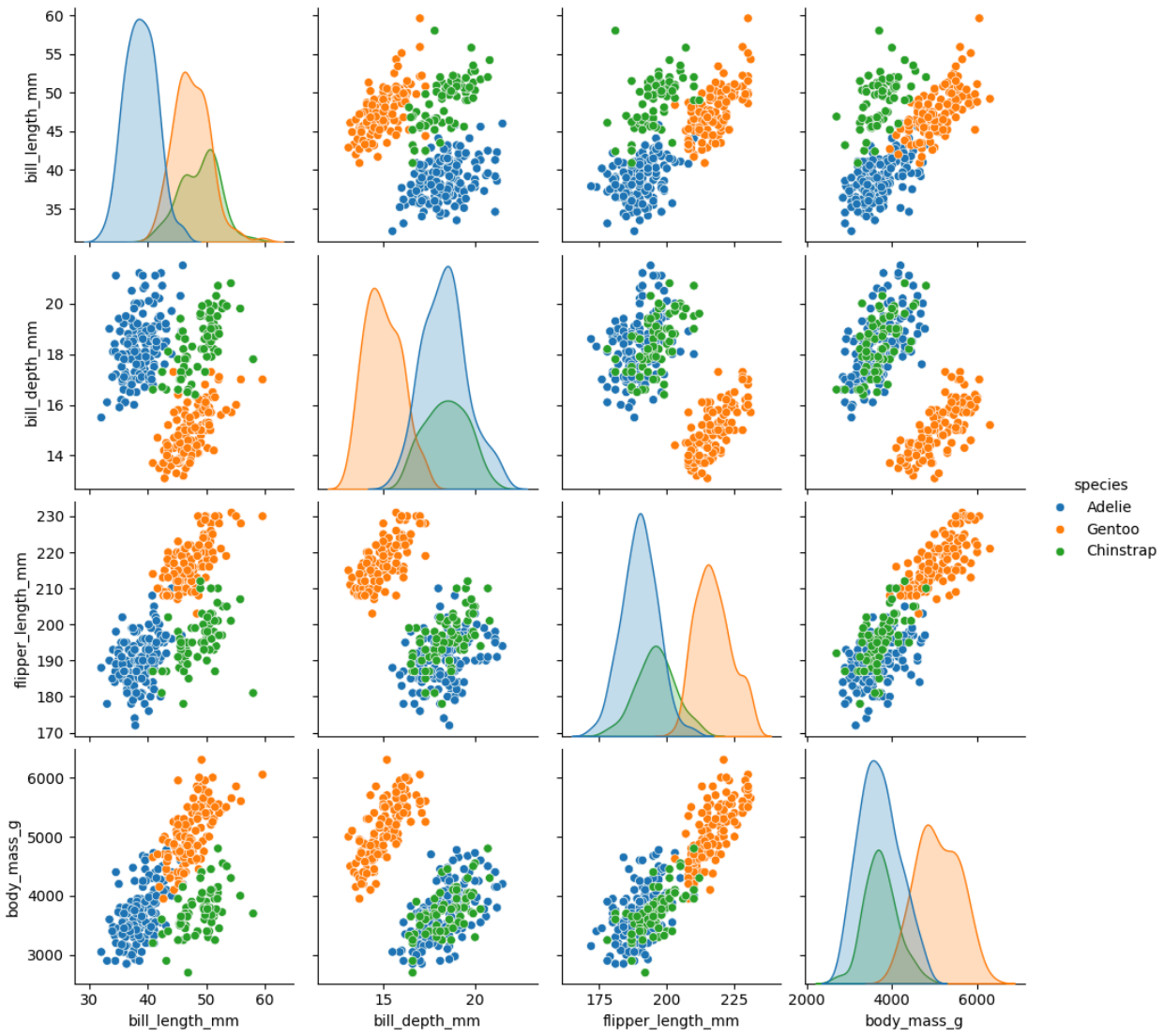
In [30]:
```
plt.scatter(a_adelie['bill_length_mm'],a_adelie['bill_depth_mm'],label='Adelie
plt.scatter(a_gentoo['bill_length_mm'],a_gentoo['bill_depth_mm'],label='Gentoo
plt.scatter(a_chinstrap['bill_length_mm'],a_chinstrap['bill_depth_mm'],label='
plt.xlabel('bill_length_mm')
plt.ylabel('bill_depth_mm')
plt.legend()
```

Out[30]:
```
<matplotlib.legend.Legend at 0x2e6143b02f0>
```



In [35]:
```
sns.pairplot(a,hue='species')
```

Out[35]: <seaborn.axisgrid.PairGrid at 0x2e612077360>


bill_length_mm
bill_depth_mm
flipper_length_mm
body_mass_g
species
Adelie
Gentoo
Chinstrap

In [ ]: