```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

```
from google.colab import files
import pandas as pd
uploaded = files.upload()
df = pd.read_csv("pen.csv")
df.head()
```

Choose Files pen.csv

pen.csv(text/csv) - 13862 bytes, last modified: 9/27/2025 - 100% done

Saving pen.csv to pen.csv

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

Next steps: Generate code with df | New interactive sheet

```
df = pd.read_csv('pen.csv')
df
```

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

344 rows × 7 columns

Next steps: Generate code with df | New interactive sheet

```
# Separate by species
df_Adelie = df[df['species'] == 'Adelie']
df_Gentoo = df[df['species'] == 'Gentoo']
df_Chinstrap = df[df['species'] == 'Chinstrap']
```

```
df_Adelie
```

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

152 rows × 7 columns

Next steps: Generate code with df_Adelie | New interactive sheet

```
df_Adelie.shape
```

```
(152, 7)
```

```
df_Gentoo
```

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

124 rows × 7 columns

Next steps: Generate code with df_Gentoo | New interactive sheet

```
df_Gentoo.shape
```

```
(124, 7)
```

```
df_Chinstrap
```

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

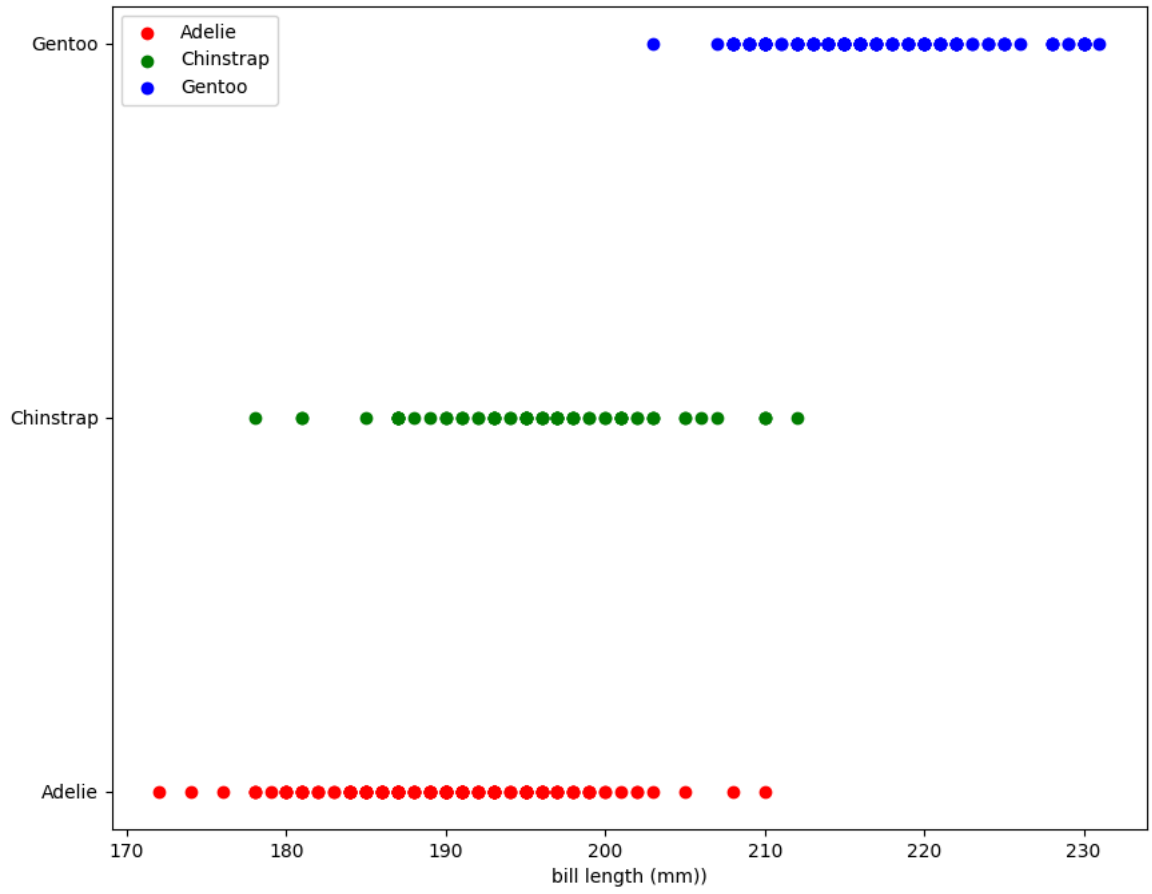
68 rows × 7 columns

Next steps: Generate code with df_Chinstrap | New interactive sheet

```
df_Chinstrap.shape
```

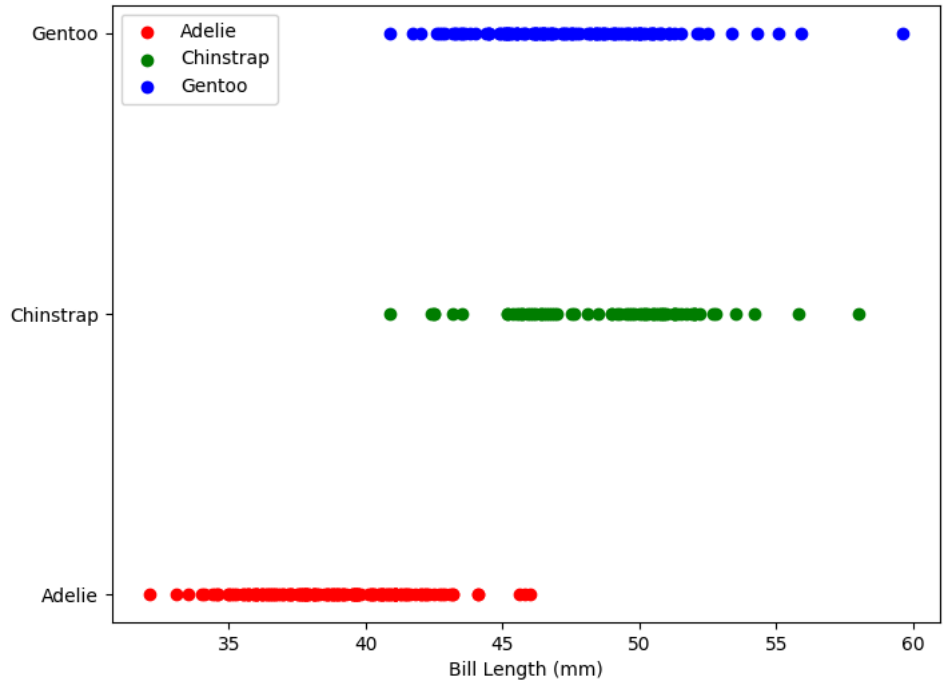
```
(68, 7)
```

```
# Scatter plot (with offsets so points don't overlap)
plt.figure(figsize=(10, 8))
plt.scatter(df_Adelie['flipper_length_mm'], np.zeros_like(df_Adelie['flipper_length_mm']),
            label='Adelie', color='red')
plt.scatter(df_Chinstrap['flipper_length_mm'], np.ones_like(df_Chinstrap['flipper_length_mm']),
            label='Chinstrap', color='green')
plt.scatter(df_Gentoo['flipper_length_mm'], np.full_like(df_Gentoo['flipper_length_mm'], 2),
            label='Gentoo', color='blue')

plt.xlabel('bill length (mm))')
plt.yticks([0, 1, 2], ['Adelie', 'Chinstrap', 'Gentoo'])
plt.legend()
plt.show()
```

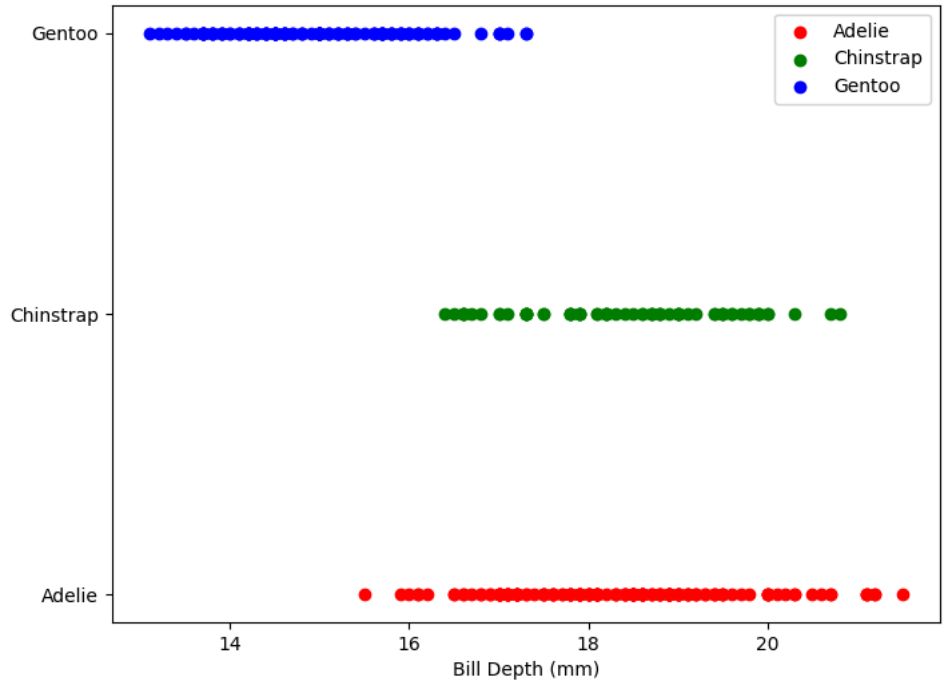


Start coding or generate with AI.

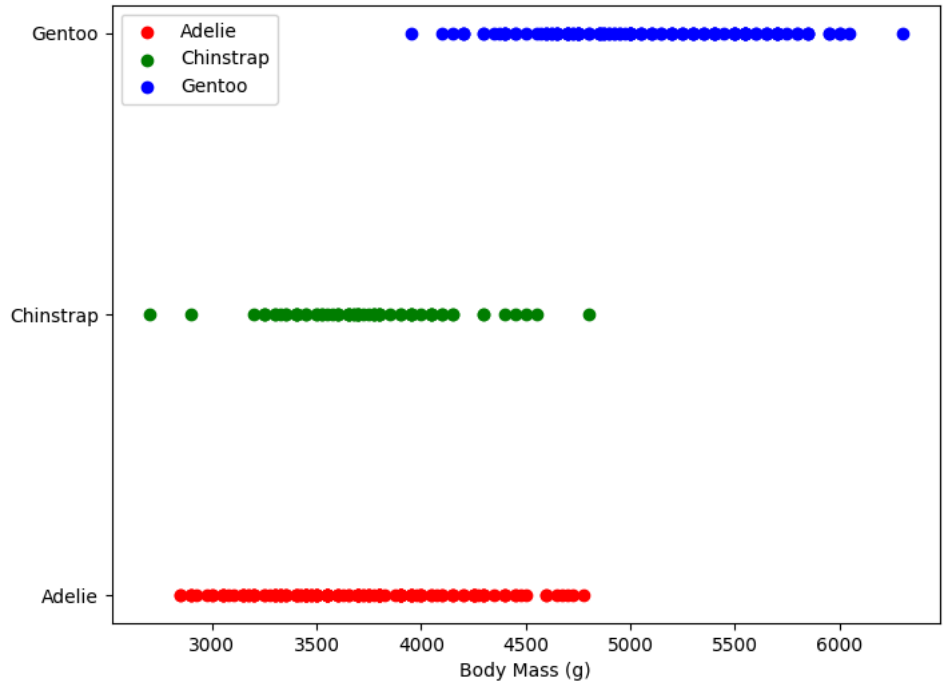
```
plt.figure(figsize=(8, 6))
plt.scatter(df_Adelie['bill_length_mm'], np.zeros_like(df_Adelie['bill_length_mm']),
            label='Adelie', color='red')
plt.scatter(df_Chinstrap['bill_length_mm'], np.ones_like(df_Chinstrap['bill_length_mm']),
            label='Chinstrap', color='green')
plt.scatter(df_Gentoo['bill_length_mm'], np.full_like(df_Gentoo['bill_length_mm'], 2),
            label='Gentoo', color='blue')

plt.xlabel('Bill Length (mm)')
plt.yticks([0, 1, 2], ['Adelie', 'Chinstrap', 'Gentoo'])
plt.legend()
plt.show()
```




```
plt.figure(figsize=(8, 6))
plt.scatter(df_Adelie['bill_depth_mm'], np.zeros_like(df_Adelie['bill_depth_mm']),
            label='Adelie', color='red')
plt.scatter(df_Chinstrap['bill_depth_mm'], np.ones_like(df_Chinstrap['bill_depth_mm']),
            label='Chinstrap', color='green')
plt.scatter(df_Gentoo['bill_depth_mm'], np.full_like(df_Gentoo['bill_depth_mm'], 2),
            label='Gentoo', color='blue')

plt.xlabel('Bill Depth (mm)')
plt.yticks([0, 1, 2], ['Adelie', 'Chinstrap', 'Gentoo'])
plt.legend()
plt.show()
```



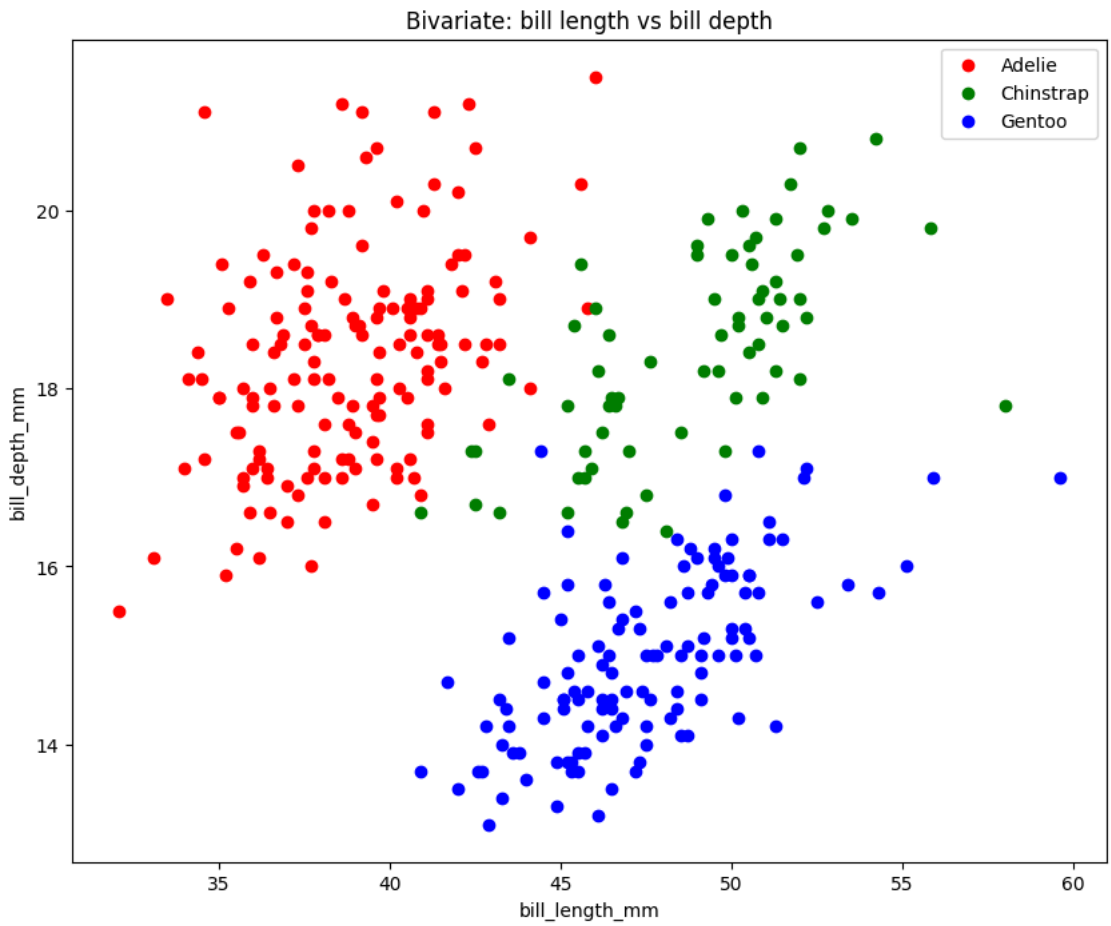

```
plt.figure(figsize=(8, 6))
plt.scatter(df_Adelie['body_mass_g'], np.zeros_like(df_Adelie['body_mass_g']),
            label='Adelie', color='red')
plt.scatter(df_Chinstrap['body_mass_g'], np.ones_like(df_Chinstrap['body_mass_g']),
            label='Chinstrap', color='green')
plt.scatter(df_Gentoo['body_mass_g'], np.full_like(df_Gentoo['body_mass_g'], 2),
            label='Gentoo', color='blue')

plt.xlabel('Body Mass (g)')
plt.yticks([0, 1, 2], ['Adelie', 'Chinstrap', 'Gentoo'])
plt.legend()
plt.show()
```



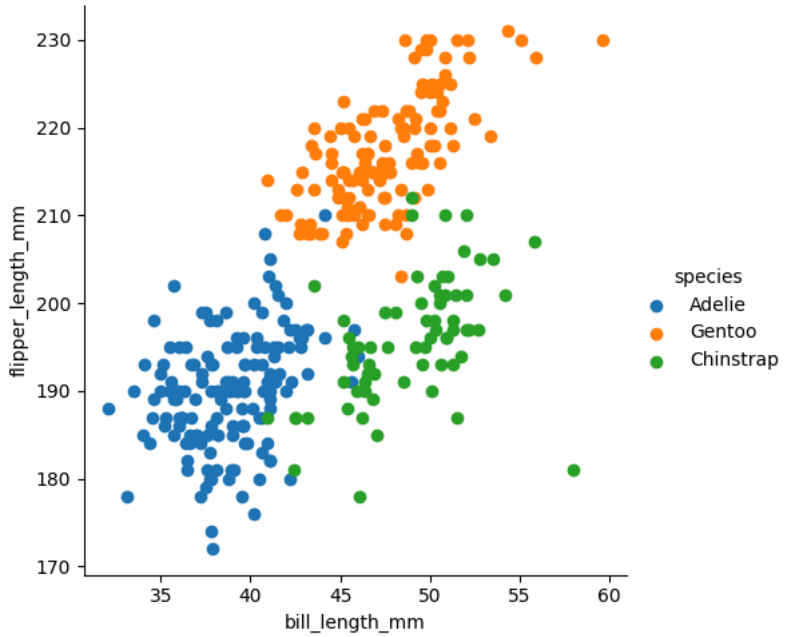
Double-click (or enter) to edit

```
# Bivariate scatter plot
plt.figure(figsize=(10, 8))
plt.scatter(df_Adelie['bill_length_mm'], df_Adelie['bill_depth_mm'], color='red', label='Adelie')
plt.scatter(df_Chinstrap['bill_length_mm'], df_Chinstrap['bill_depth_mm'], color='green', label='Chinstrap')
plt.scatter(df_Gentoo['bill_length_mm'], df_Gentoo['bill_depth_mm'], color='blue', label='Gentoo')

plt.xlabel('bill_length_mm')
```

```
plt.ylabel('bill_depth_mm')
plt.title('Bivariate: bill length vs bill depth')
plt.legend()
plt.show()
```




```
sns.FacetGrid(df, hue="species", height=5) \
   .map(plt.scatter, "bill_length_mm", "flipper_length_mm") \
   .add_legend()

plt.show()
```




```
import seaborn as sns

sns.pairplot(df, hue='species', diag_kind='kde', palette='Set1')
```

```
plt.suptitle('Pairplot: Multivariate Relationships (Penguins features)', y=1.02)
plt.show()
```