Bivariate Analysis

Correlations

In [2]:
```
import pandas as pd
import seaborn as sns
```

In [3]:
```
penguins = sns.load_dataset("penguins")
```

In [4]:
```
penguins.head()
```

Out[4]:

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

In [5]:
```
dataFrame = pd.DataFrame({'bill_length_mm':penguins['bill_length_mm'], 'bill_depth_mm':p
```

In [6]:
```
correlations = dataFrame.corr()
correlations.style.background_gradient(cmap='coolwarm', axis=None).format(precision=2)
```

Out[6]:

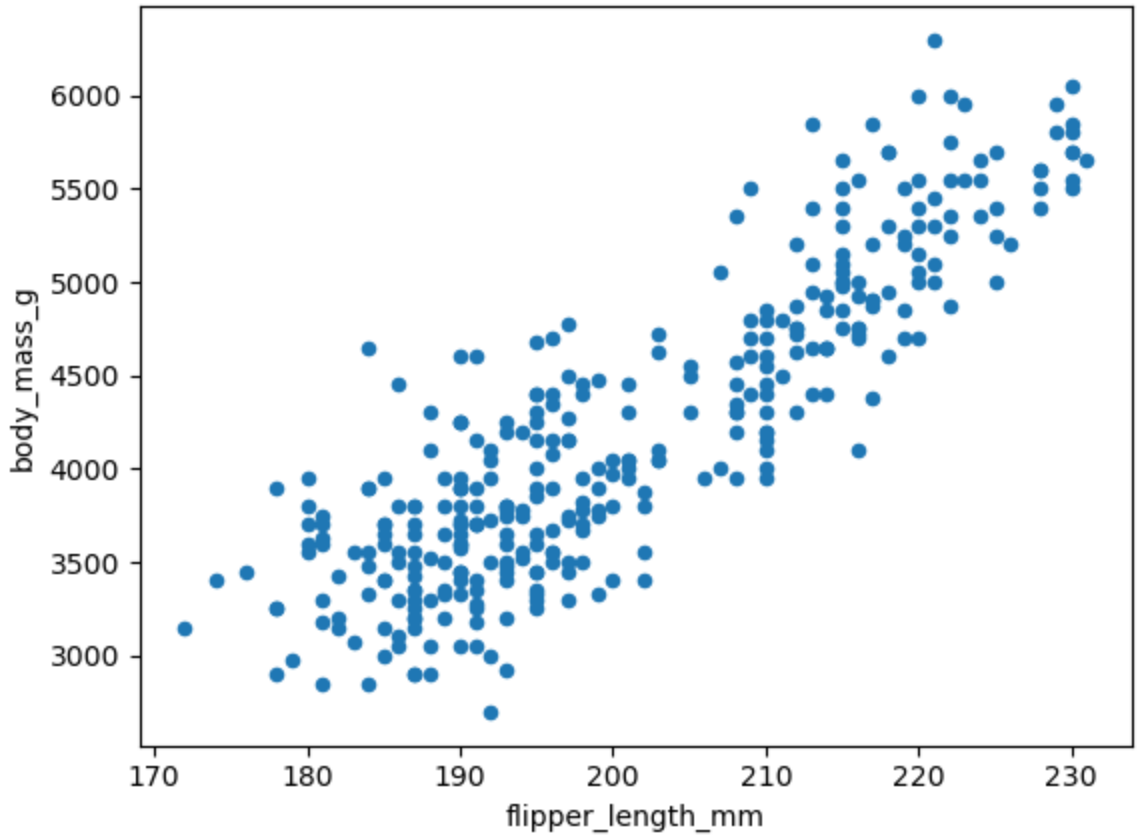
	bill_length_mm	bill_depth_mm	flipper_length_mm	body_mass_g
bill_length_mm	1.00	-0.24	0.66	0.60
bill_depth_mm	-0.24	1.00	-0.58	-0.47
flipper_length_mm	0.66	-0.58	1.00	0.87
body_mass_g	0.60	-0.47	0.87	1.00

Bij de correlations is te zien dat de sterkste correlatie tussen flipper length en body mass ligt, wat zou kunnen liggen aan dat penguins die groter zijn waarschijnlijk ook grotere flippers hebben. Hoewel bill length en bill depth beide met de snavel te maken hebben, lijkt er weinig tot geen correlatie te zijn tussen de lengte en de diepte hiervan.

Strongest positive correlation

In [7]:
```
penguins.plot(kind='scatter', x='flipper_length_mm', y='body_mass_g')
```

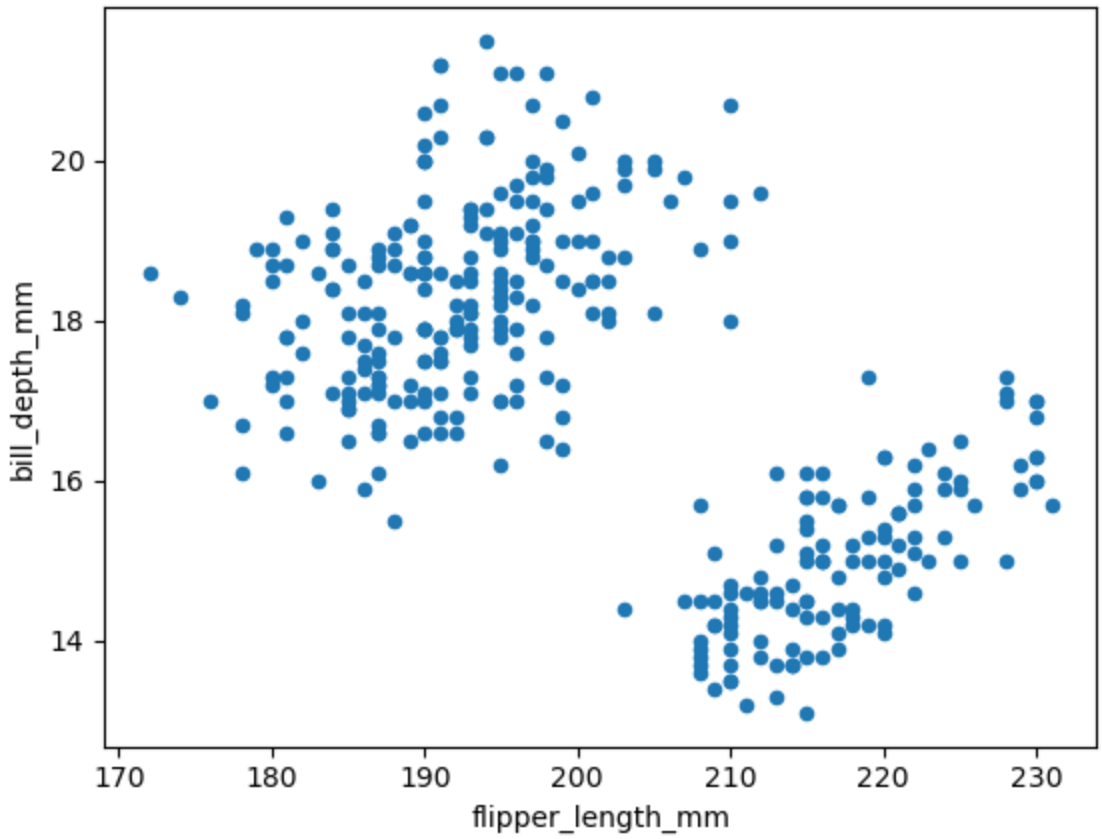
Out[7]:
```
<AxesSubplot:xlabel='flipper_length_mm', ylabel='body_mass_g'>
```



Strongest negative correlation

In [8]:
```
penguins.plot(kind='scatter', x='flipper_length_mm', y='bill_depth_mm')
```

Out[8]:
```
<AxesSubplot:xlabel='flipper_length_mm', ylabel='bill_depth_mm'>
```



Weakest correlation

In [9]:
```
penguins.plot(kind='scatter', x='bill_length_mm', y='bill_depth_mm')
```

Out[9]:
```
<AxesSubplot:xlabel='bill_length_mm', ylabel='bill_depth_mm'>
```