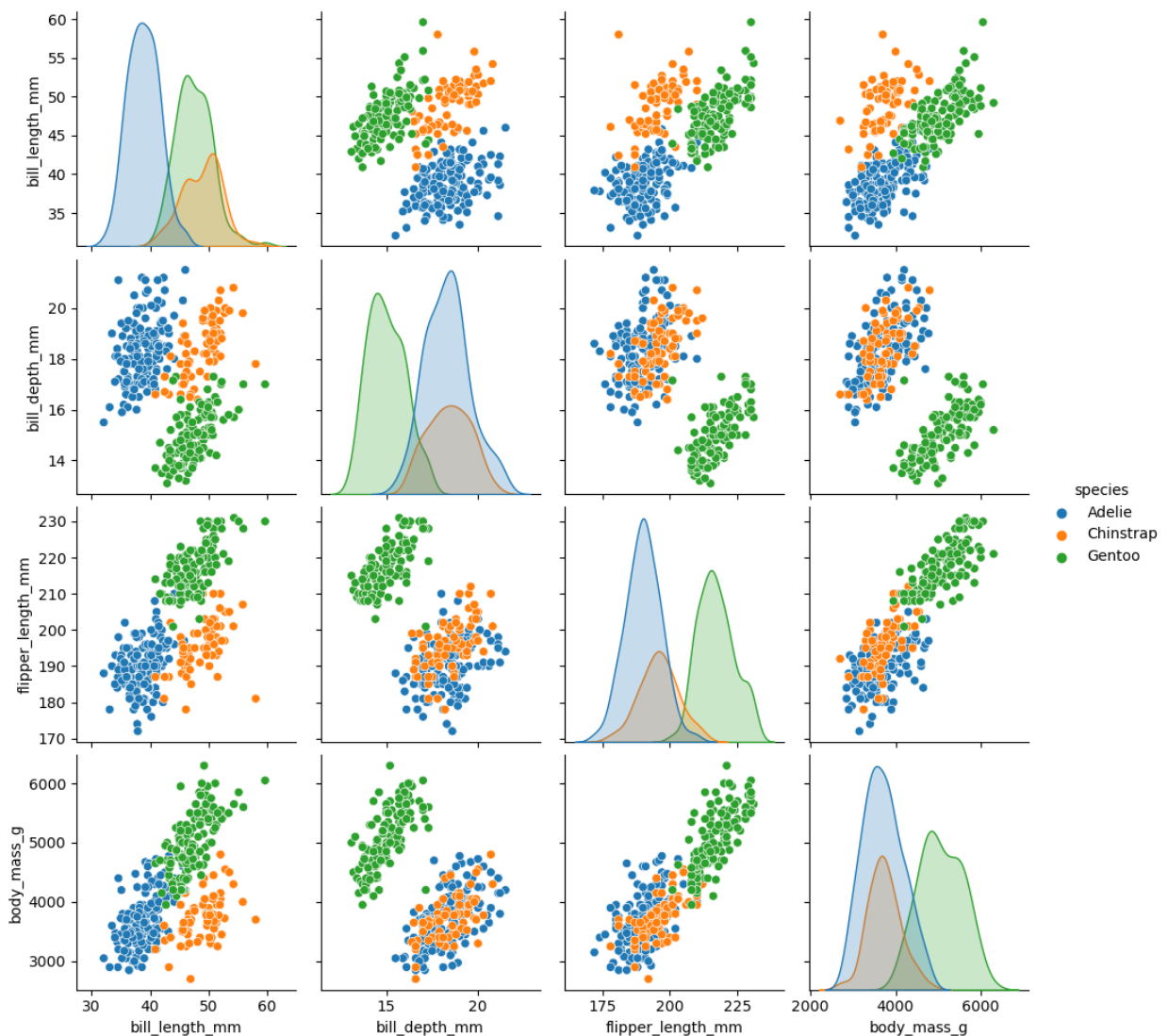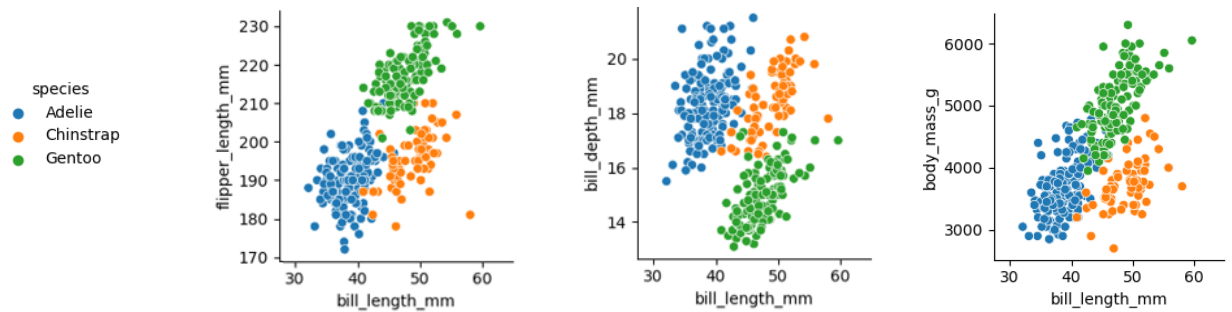# Python Ex4
Danit Noa Yechezkel
Keren Halpert

TASK 1:

1. As depicted in this pair-plot, it is fairly easy to see that one of the best features is the "bill_length_mm" feature, the reason is that the three groups of penguin species are most segregated using this classifier as opposed to other features by comparison.

Now that the first feature was chosen, we need to decide which will be the second feature for the GNB 2-feature classifier.
In order to do so, we will look for the plot that is most segregated by color between the three plots using the first, chosen feature.

It is easy to see that the plot using the "body_mass_g" is not as good as the other two.
But it is harder to tell between "flipper_length_mm" and "bill_depth_mm".
So for this reason we will check both out using GNB and the accuracy score.

The accuracy score for the pair: 'bill_length_mm' and 'bill_depth_mm' is
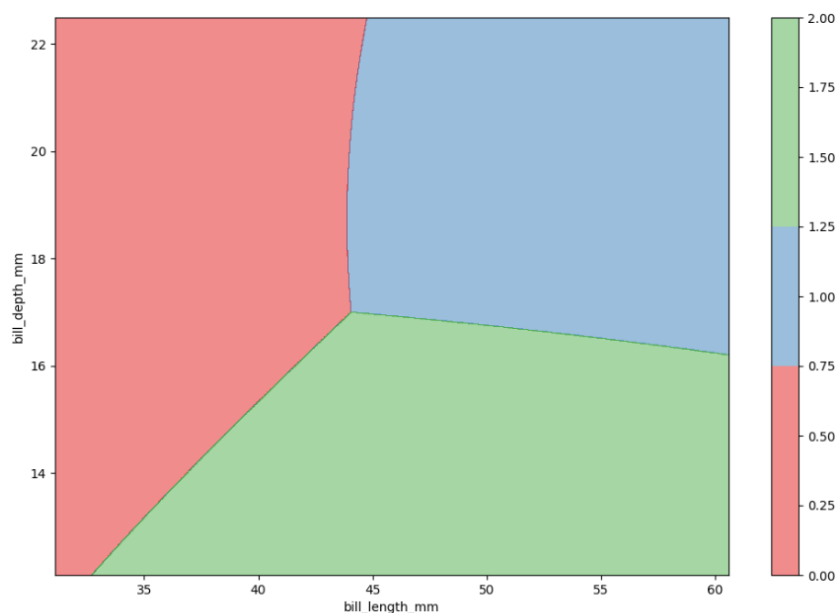
```
0.9420289855072463
```

The accuracy score for the pair: 'bill_length_mm' and 'flipper_length_mm' is
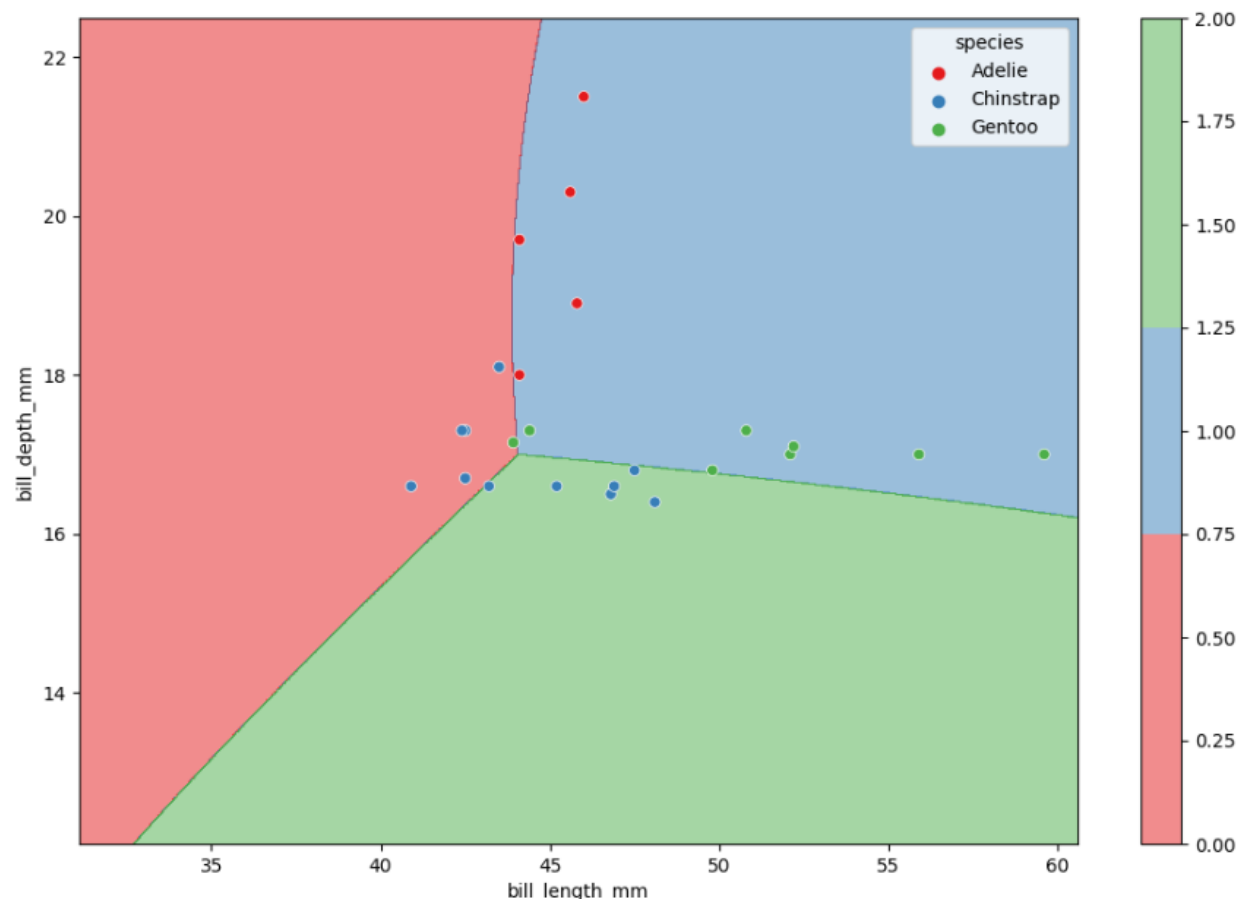
```
0.927536231884058
```

Now we can see without a doubt that the winner of the two is 'bill_depth_mm'.

3.   Filled contour plot:

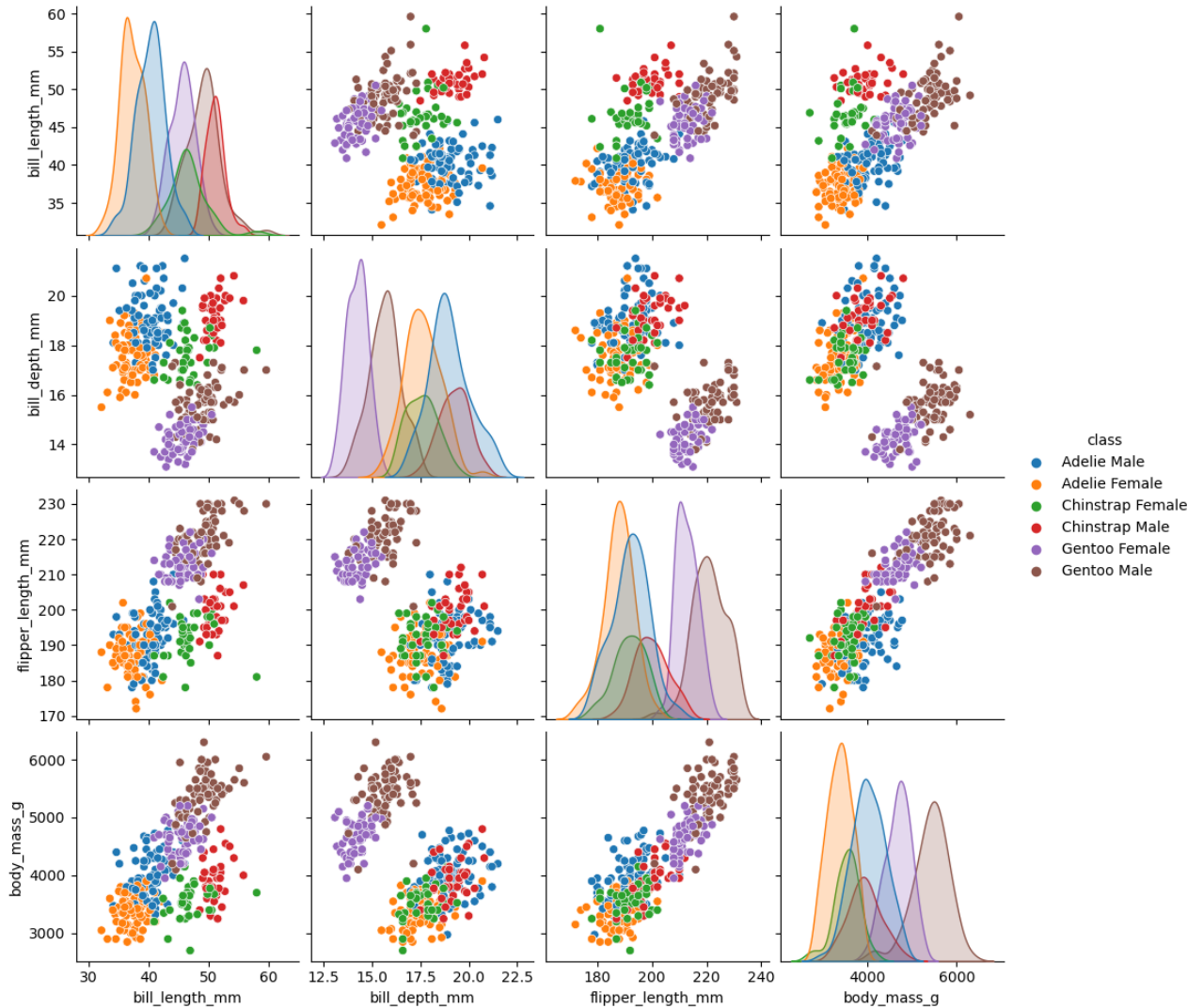4.　　Filled contour plot with scatter plot overlay of failed predictions:

5.　　Classification report:

	precision	recall	f1-score	support
Adelie	1.00	0.94	0.97	34
Chinstrap	0.82	0.82	0.82	11
Gentoo	0.92	1.00	0.96	24
accuracy			0.94	69
macro avg	0.91	0.92	0.92	69
weighted avg	0.94	0.94	0.94	69

TASK 2:

1. As done in task 1, we need to find the pair of features that provide the most segregated distribution of points.
   Again, the first choice is fairly simple: "bill_length_mm".

But the second choice is a close call visually, therefore we will check out the GNB accuracy score.

The accuracy score for the pair: 'bill_length_mm' and 'bill_depth_mm' is

```
0.8695652173913043
```

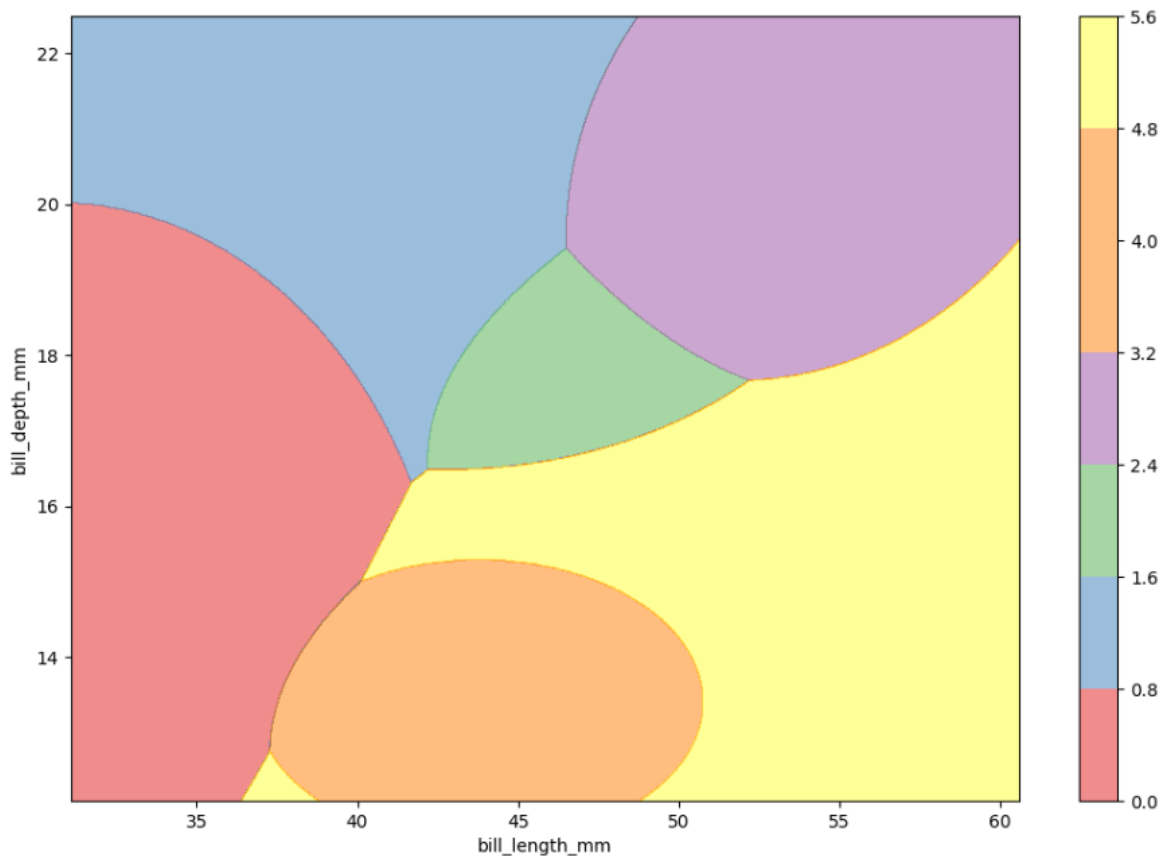The accuracy score for the pair: 'bill_length_mm' and 'flipper_length_mm' is

```
0.7826086956521740
```

The accuracy score for the pair: 'bill_length_mm' and 'body_mass_g' is
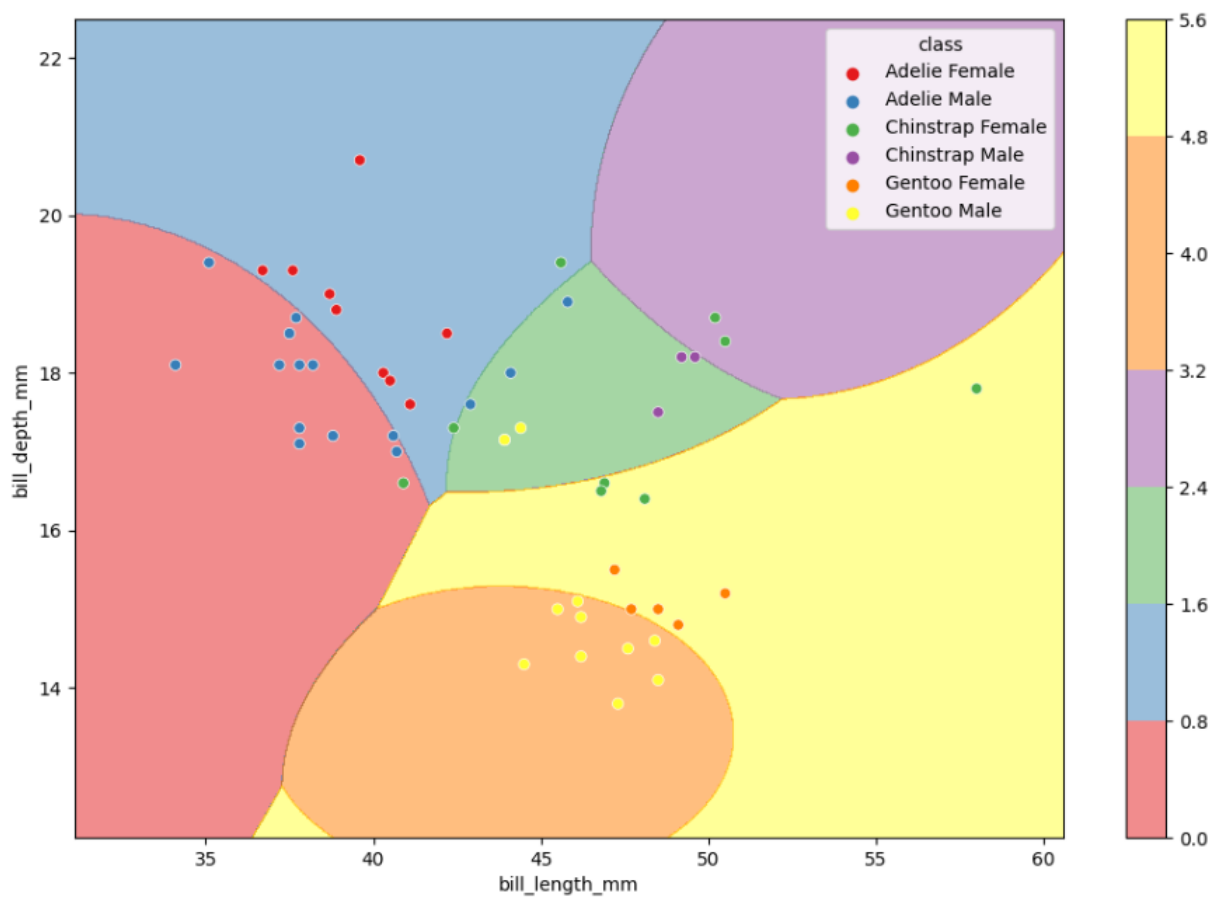
```
0.8260869565217391
```

The clear winner is, again, 'bill_depth_mm' .

3. Filled contour plot:

4. Filled contour plot with scatter plot overlay of failed predictions:

5. Classification report:

	precision	recall	f1-score	support
Adelie Female	0.92	0.75	0.83	16
Adelie Male	0.76	0.89	0.82	18
Chinstrap Female	0.50	0.50	0.50	4
Chinstrap Male	1.00	0.86	0.92	7
Gentoo Female	1.00	1.00	1.00	7
Gentoo Male	0.94	1.00	0.97	17
accuracy			0.87	69
macro avg	0.85	0.83	0.84	69
weighted avg	0.88	0.87	0.87	69