Exercise Set 2

This report contains my answers for exercise set 2.

Problem 8

Task a

Coefficients for penguins train data fit.

```
    (Intercept)     bill_length_mm      bill_depth_mm flipper_length_mm
  -380.81882000        14.30428414        -11.68758439        0.48138045
    body_mass_g
    -0.03094901
```

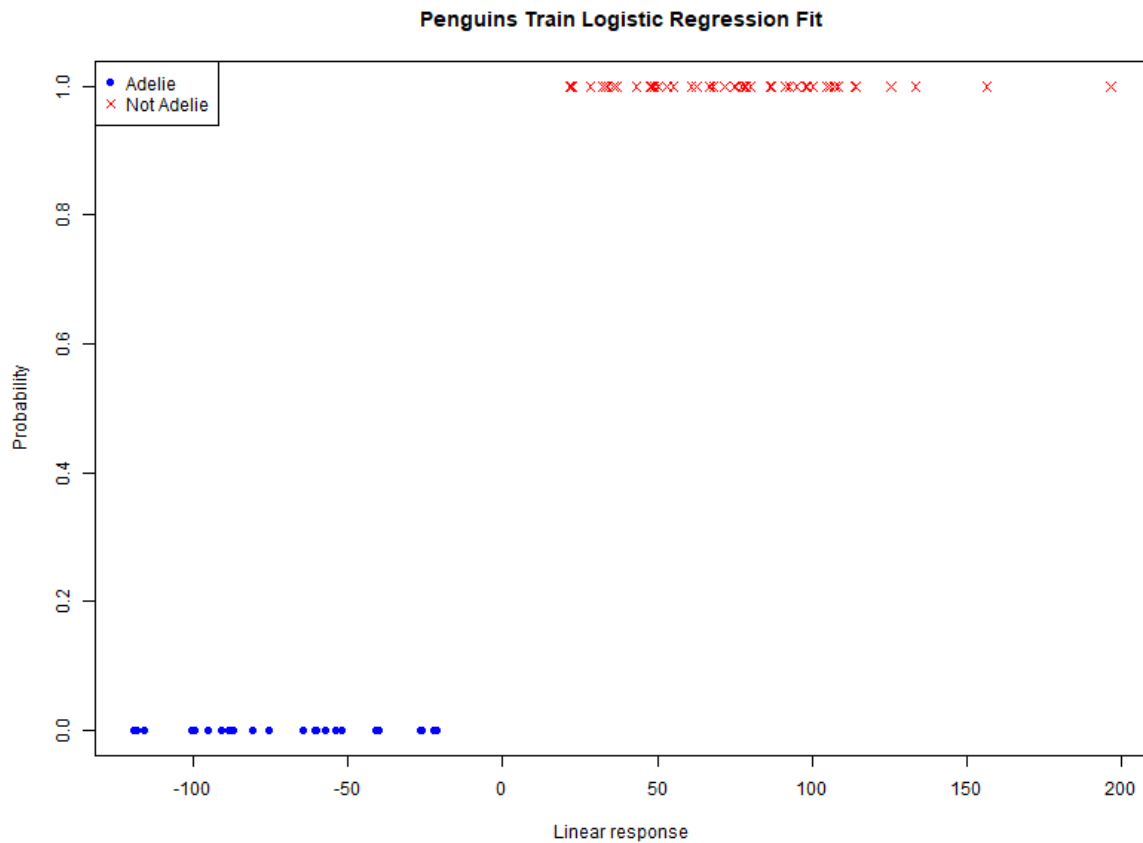Accuracy on training and test sets without regularisation:

```
[1] "Train accuracy:"
```

```
[1] 1
```

```
[1] "Test accuracy:"
```

```
[1] 0.9466667
```

Plot of penguins train data predictions.

Penguins Train Logistic Regression Fit

Task b

Coefficients for penguins train data fit.

```
Loading required package: Matrix

5 x 1 sparse Matrix of class "dgCMatrix"
                              s0
(Intercept)        -16.082633193
bill_length_mm       0.662337141
bill_depth_mm       -0.733855638
flipper_length_mm    0.004030675
body_mass_g              .
```

Accuracy on training and test sets using lasso with regularisation:

```
[1] "Train accuracy:"

[1] 1

[1] "Test accuracy:"

[1] 0.9866667
```

Task c

R gives warnings because the model is fitting the data too well and because there's no regulation, it will try to push probabilities closer to 0 or 1 forever if the programme didn't stop itself and give the warning.