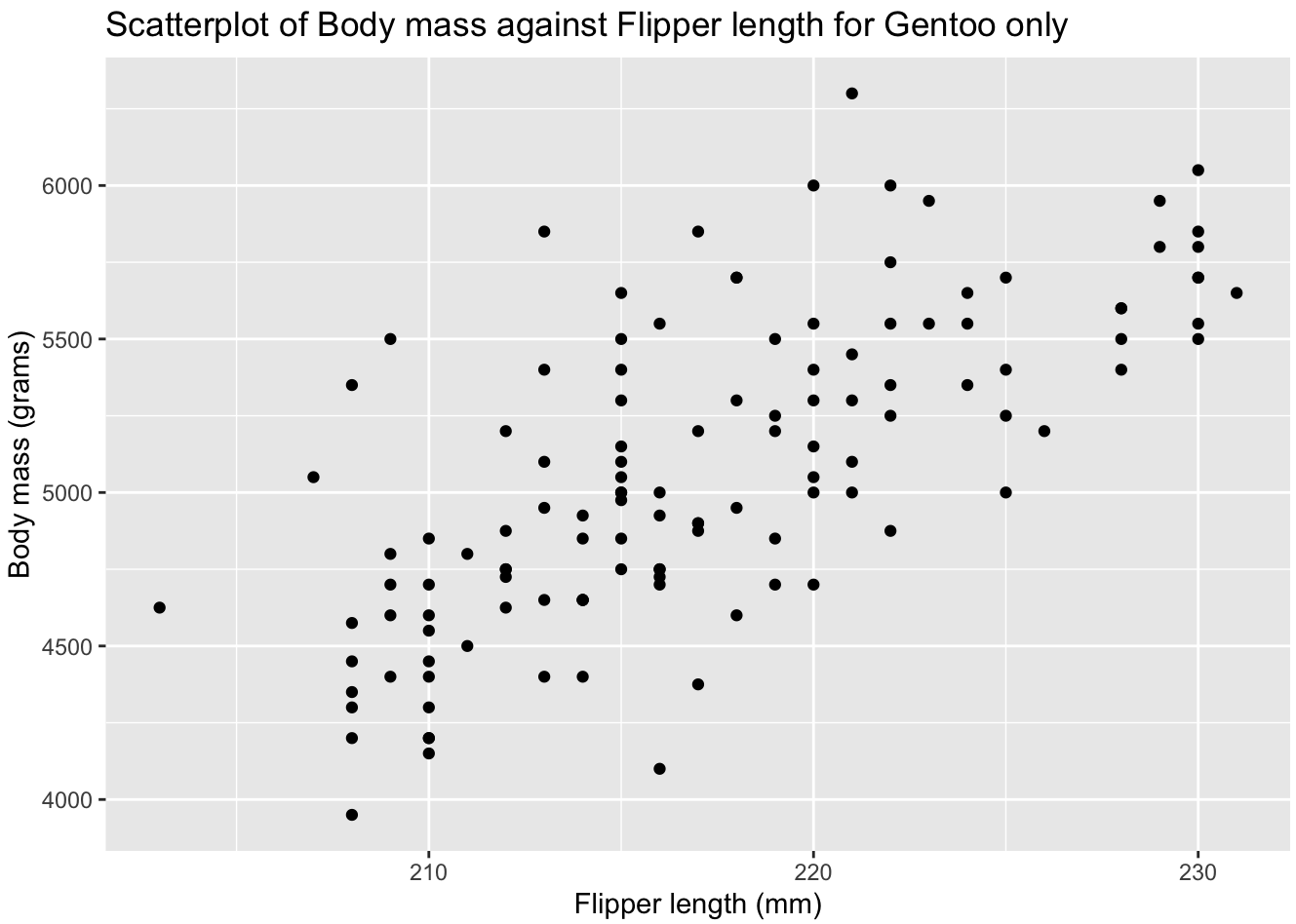
M4_examples

1. Produce a plot of body mass against flipper length for Gentoo penguins. Write the estimated linear regression equation.

```
ggplot(penguins[which(penguins$species=='Gentoo'), ], aes(x=flipper_length_mm, y=body_mass_g))+
  geom_point()+
  labs(x="Flipper length (mm)", y="Body mass (grams)",
       title="Scatterplot of Body mass against Flipper length for Gentoo only")
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```



```
Gentoo <- na.omit(penguins[which(penguins$species=='Gentoo'),
                  c("flipper_length_mm", "body_mass_g")])

result <- lm(body_mass_g ~ flipper_length_mm, data=Gentoo)
summary(result)
```

```
##
## Call:
## lm(formula = body_mass_g ~ flipper_length_mm, data = Gentoo)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -911.18 -235.76  -51.93  170.75 1015.71
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       -6787.281   1092.552  -6.212 7.65e-09 ***
## flipper_length_mm    54.623      5.028  10.863  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 360.2 on 121 degrees of freedom
## Multiple R-squared:  0.4937, Adjusted R-squared:  0.4896
## F-statistic:   118 on 1 and 121 DF,  p-value: < 2.2e-16
```

Note: body_mass_g = -6787.281 + 54 * (flipper_length_mm)

2. What is the change in the predicted body mass (in grams) when flipper length increases by 1mm, for Gentoo penguins? Also report the corresponding 95% confidence interval for the change in the predicted body mass (in grams) when flipper length increases by 1mm.

Note: the change in y_hat is 54.623 grams. 95% for the change is given by B1/se(B1):

```
confint(result, level = 0.95)
```

```
##                        2.5 %      97.5 %
## (Intercept)       -8950.27535 -4624.28587
## flipper_length_mm    44.66777    64.57724
```

3. Conduct a hypothesis test to determine whether or not there is a linear association between body mass and flipper length for Gentoo penguins. State the hypotheses, p-value, and conclusion in context.

Note: H0: B1 = 0 and Ha: B1 != 0. With p-value for F-statistic of 2.2e-16 we have to reject H0 and conclude that there is not enough evidence to state that the slope is zero and thus there is no relationship mass ~ length.

4. Are your results from parts 2 and 3 consistent? Briefly explain.

Note: yes, there are, because from #2 it appears that B1 is in 95% CI and from #3 it appears that F-test is highly significants as well.

5. Estimate the mean body mass (in grams) for Gentoo penguins with flipper lengths of 200mm. Also report the 95% confidence interval for the mean body mass (in grams) for Gentoo penguins with flipper lengths of 200mm.

```
-6787.281 + 54 * (200)
```

```
## [1] 4012.719
```

```
new_data <- data.frame(flipper_length_mm=200)
predict(result, new_data, level=0.95, interval="confidence")
```

```
##       fit      lwr      upr
## 1 4137.22 3954.446 4319.993
```

6. Report the 95% prediction interval for the body mass (in grams) of a Gentoo penguin with flipper length of 200mm.

```
new_data <- data.frame(flipper_length_mm=200)
predict(result, new_data, level=0.95, interval="prediction")
```

```
##       fit      lwr      upr
## 1 4137.22 3401.121 4873.319
```

7. A researcher hypothesizes that for Gentoo penguins, the predicted body mass increases by more than 50 g for each additional mm in flipper length. Conduct an appropriate hypothesis test. What is the null and alternative hypotheses, test statistic, and conclusion?

Note ???

- H0: flipper_length > 50
- Ha: flipper_length =< 50
- test statistic is given by the following

```
pt((50-54.623) / 5.028, 121)
```

```
## [1] 0.1798445
```

```
confint(result, level = 0.95)
```

```
##                         2.5 %      97.5 %
## (Intercept)       -8950.27535 -4624.28587
## flipper_length_mm    44.66777    64.57724
```