Guided Question Set 4 Solutions

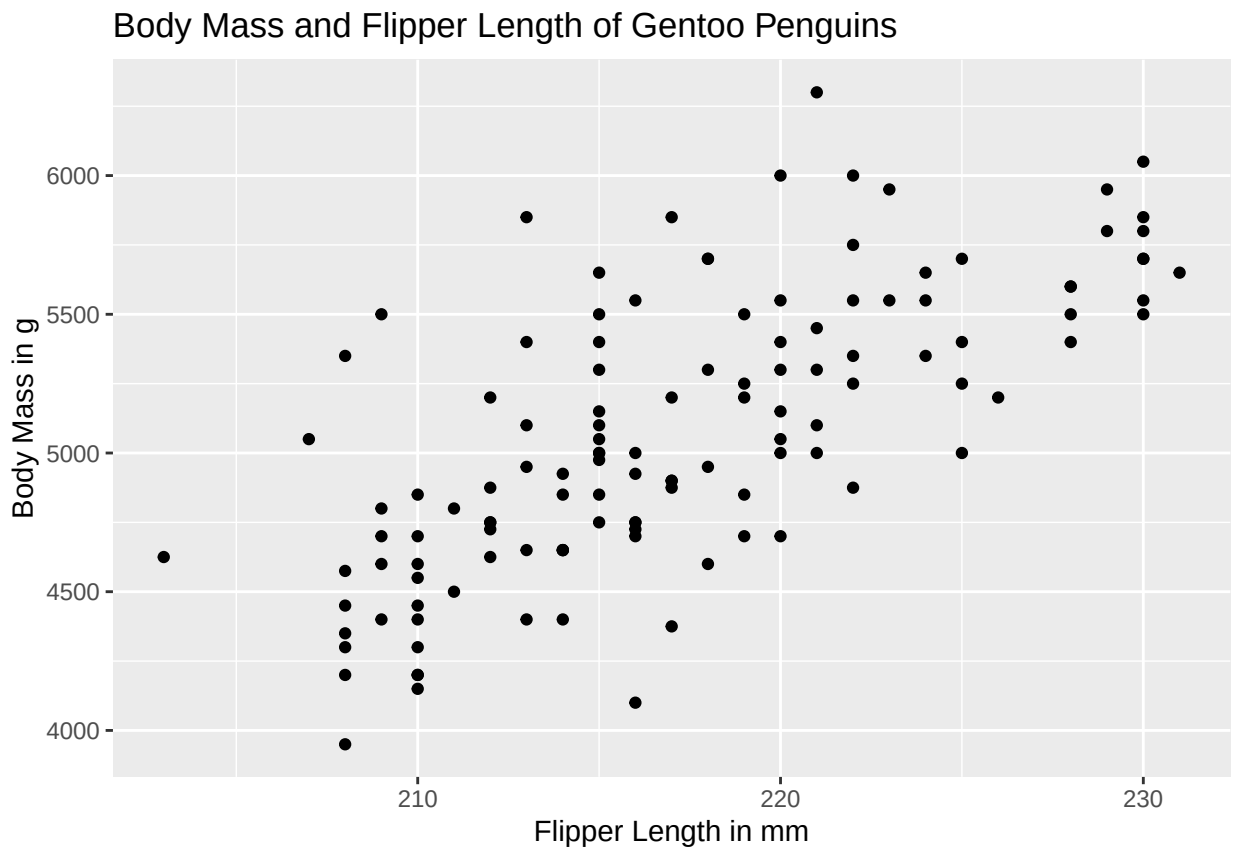
```
library(tidyverse)
library(palmerpenguins)
Data<-penguins
```

1)

```
gentoo<-Data %>%
  filter(species=="Gentoo")

ggplot(gentoo, aes(x=flipper_length_mm,y=body_mass_g))+
  geom_point()+
  labs(x="Flipper Length in mm",
       y="Body Mass in g",
       title="Body Mass and Flipper Length of Gentoo Penguins")
```

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



The scatterplot of body mass against flipper length for Gentoo penguins is displayed above.

```
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
```

```
##   (1 observation deleted due to missingness)
## Multiple R-squared:  0.4937, Adjusted R-squared:  0.4896
## F-statistic:   118 on 1 and 121 DF,  p-value: < 2.2e-16
```

The estimated regression equation is $\hat{y} = -6787.281 + 54.623x$.

2)

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

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

The predicted body mass increases by 54.623g per mm increase in flipper length, for Gentoo penguins. The corresponding 95% confidence interval is (44.67, 64.58) g.

3)

$H_0 : \beta_1 = 0, H_a : \beta_1 \neq 0.$

The t statistic is 10.863. Since the corresponding p-value is less than 0.05, we reject the null hypothesis. The data support the claim that there is a linear relationship between body mass and flipper length for Gentoo penguins.

4)

Yes the results are consistent. The 95% CI excluded the value of 0 (the value under the null hypothesis), and we rejected the null hypothesis at 0.05 significance level.

5)

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

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

The mean body mass is 4137.22g for Gentoo penguins with flipper length 200mm. The corresponding 95% confidence interval is (3954.446, 4319.993)g.

6)

```
predict(result, newdata,
        level=0.95, interval = "prediction")
```

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

The 95% prediction interval for the body mass of a Gentoo penguin with flipper length 200mm is (3401.121, 4873.319)g.

7)

$H_0 : \beta_1 = 50, H_a : \beta_1 > 50.$

The t stat is $t = \frac{54.623-50}{5.028} = 0.919$.

The pvalue is `1 - pt(0.919, 121)` which is 0.1798.

The critical value is `qt(0.95,121)` which is 1.658.

So we fail to reject the null hypothesis. The data do not support the researcher's claim that for Gentoo penguins, the predicted body mass increases by more than 50 g for each additional mm in flipper length.