Chloe Lang
Professor Nelson
Lab 6 Report
October 24, 2021

Q1

```
rm(list = ls())
sse_mean = function (x, na.rm = TRUE) sd (x, na.rm = TRUE) / sqrt(length(x))
sse_mean(penguins$body_mass_g)
sse_mean(mtcars$mpg)
```

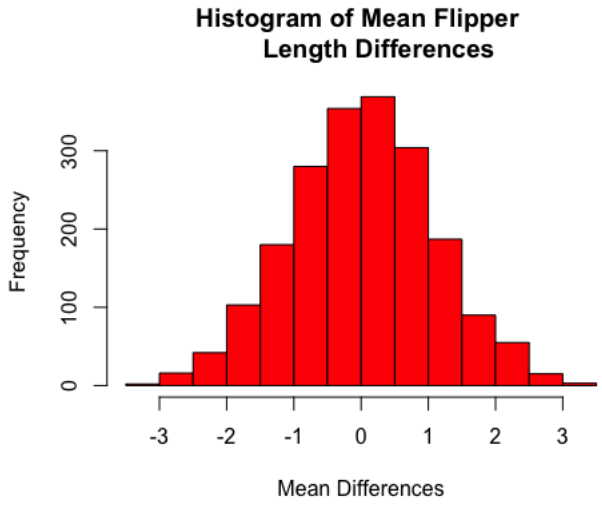
Q2

```
two_group_resample = function(x, n_1, n_2)
{
  dat_pen = droplevels(subset(penguins, species != "Gentoo"))
  x = dat_pen$flipper_length_mm
  n_1 = 68
  n_2 = 152

  dat_1 = sample(x, n_1, replace = TRUE)
  dat_2 = sample(x, n_2, replace = TRUE)

  difference_in_means = mean(dat_1, na.rm = TRUE) - mean(dat_2, na.rm = TRUE)
  difference_in_means
}
```

Q3

My function performs Monte Carlo i.e., a null hypothesis.

Q4

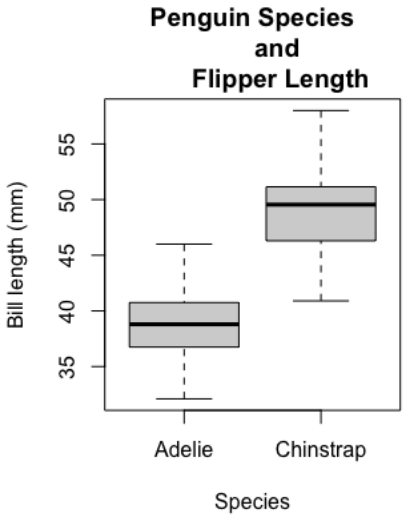


Q5 There are zero resampled differences of means greater than 5.8

```
agg_means = aggregate(
  flipper_length_mm ~ species,
  data = dat_pen,
  FUN = "mean",
  na.rm = TRUE)
diff_observed = diff(agg_means[,2])
sum(abs(mean_differences) >= diff_observed)
```

Q6 Given that your p-value is less than 1 per 10 million, I would expect to do 10 million simulations at a minimum to see a difference in mean flipper length greater than 5.8 mm.

Q7



Q8

```
agg_pen_means = aggregate(
  bill_length_mm ~ species,
  data = dat_pen,
  FUN = mean,
  na.rm = TRUE)
```

Group means:

```
    species bill_length_mm
1    Adelie       38.79139
2 Chinstrap       48.83382
```

```
diff_crit = diff(agg_pen_means[,2])
```

Difference in means: 10.04243

Q9 The p-value is 2.2e-16 which is giving a very small p-value. This indicates that there is a significant difference in values between the two species based on the bill length. With this, we can likely reject the null hypothesis, given that there is a significant difference in values between the two variables.

Q10 There were zero differences in means greater than the diff_crit.

Q11