Ian Eggleston

1.

```
rm(list = ls())

rope = read.csv(here("data", "rope.csv"))
rope$rope.type = factor(rope$rope.type)

n_obs = length(rope$rope.type)
n_groups = length(levels(rope$rope.type))

ss_tot = sum((rope$p.cut - mean(rope$p.cut))^2)
df_tot = n_obs - 1

agg_sum_sq_resids = aggregate(
  x = rope$p.cut,
  by = list(rope$rope.type),
  FUN = function(x) (sum((x - mean(x))^2)))
ss_within = sum(agg_sum_sq_resids$x)
df_within = n_obs - n_groups

ss_among = ss_tot - ss_within
df_among = n_groups - 1

ms_within = ss_within / (n_obs - n_groups)
ms_among  =  ss_among / (n_groups - 1)

f_ratio = ms_among / ms_within
f_pval = pf(f_ratio, df_among, df_within, lower.tail = FALSE)
```

2. Based on the boxplots, it does not seem like there is constant variance because the size of the boxes are all different.
3. bartlett.test(p.cut ~ rope.type, data = rope)
        Bartlett test of homogeneity of variances

        data: p.cut by rope.type
        Bartlett's K-squared = 19.687, df = 5, p-value = 0.00143
4. ANOVA would be appropriate for use on the raw data because there are multiple groups we are comparing. Based on the boxplots and Bartlett test, the group distributions seem to be different between one another more than within each group. Therefore, this shows that grouping them better explains the variance than looking at variance of the total group.
5. Blaze is the base rope type.
6. 0.36714

7. The XTC rope intercept is -0.10164 and base case intercept is 0.36714. Therefore, XTC mean percent of rope cut is $0.36714 - 0.10164 = 0.2655$

8. The p-value of model residuals is 7.238e-07

9. The residuals do not meet the normally assumption because we can reject the null hypothesis that the values are from a normal distribution.

10. 3 of the rope type residuals meet the normality assumption. (BS, PI, SB)

11. An ANOVA test would be appropriate here to see how effective the grouping was for the residuals. Some groups meet the normality assumption, while others do not.

Female Penguin Body Mass by Species
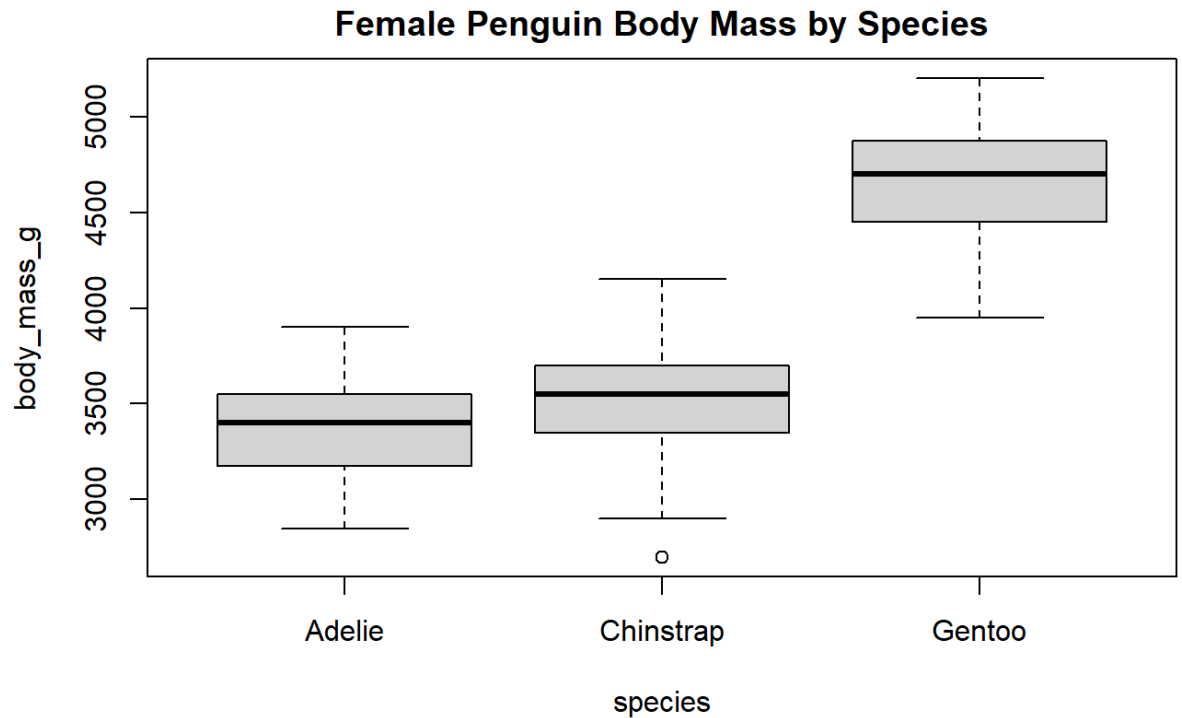
12.

13. It seems like variance may differ among species, notably the Gentoo penguins seem to have a larger variance than the other two.

14. P-value = 0.9056

   This p-value does not allow us to reject the null hypothesis for the Bartlett Test. Therefore, the homogeneity of variance assumption has been met and variance is constant across the species.

15. P-value = 0.3639

   The residual normality assumption is met because we cannot reject the null hypothesis of the Shapiro test with this p-value.

16. According to the Tukey test, all species have significantly different body masses.

17. The results match considering Gentoo is very different from the rest, the p-value was 0. There was a slight relationship between the Adelie and Chinstrap, but not enough for the p-value to be greater than 0.05, it was 0.0179.