Julia Vineyard

Worked with: Andrew Gordon, Jessica Bonin

Lab 10 ANOVA By Hand

1. require(here)

```
rope= read.csv(here("data", "rope.csv"))

rm(list = ls())

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

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

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

agg_resids= aggregate(x=rope$p.cut, by=list(rope$rope.type),
       FUN= function(x) {x-mean(x)})
str(agg_resids)

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

ss_among = ss_tot - ss_within
ss_among
df_among = n_groups - 1
df_among

ms_within =ss_within / df_within
```

```
ms_among  =ss_among / df_among

f_ratio = ms_among / ms_within
f_ratio #if a big number then reject the null that all means are the same

f_pval = pf(f_ratio, df_among, df_within, lower.tail= FALSE)
f_pval

fit_1= lm(p.cut ~ rope.type, data=rope)
anova_fit_1 =anova(fit_1)
str(anova_fit_1)

anova_fit_1$`Sum Sq`
```

2. Based on the boxplot I think that there are not equal variances among the groups.
3. p-value = 0.00143
4. No, an ANOVA analysis is not appropriate for the raw data since there is not homogeneity among the variances. This is supported by the boxplot and the p-value of the Bartlett test. The significant p-value means that we can reject the null hypothesis that the variance is the same amongst the groups.
5. BLAZE is the base case.
6. 0.36714. The base case rope is just the Estimate value of the Intercept.
7. 0.36714 - 0.10164 = 0.2655 (Intercept Estimate + rope.typeXTC Estimate)