6.4

a) aov(log(data$ADT1\_YEAST)~ ADT1\_YEAST$MIT)

#summary statistics

summary(aov(log(data$ADT1\_YEAST)~ ADT1\_YEAST$MIT))

b)To use the ANOVA test we made the following assumptions:

* Each group sample is drawn from a normally distributed population
* All populations have a common variance
* All samples are drawn independently of each other
* Within each sample, the observations are sampled randomly and independently of each other
* Factor effects are additive

The presence of outliers can also cause problems. In addition, we need to make sure that the *F* statistic is well behaved. In particular, the *F* statistic is relatively robust to violations of normality provided:

* The populations are symmetrical and uni-modal.
* The sample sizes for the groups are equal and greater than 10

In general, as long as the sample sizes are equal (called a **balanced model**) and sufficiently large, the normality assumption can be violated provided the samples are symmetrical or at least similar in shape (e.g. all are negatively skewed).

The *F* statistic is not so robust to violations of homogeneity of variances. A rule of thumb for balanced models is that if the ratio of the largest variance to smallest variance is less than 3 or 4, the F-test will be valid. If the sample sizes are unequal then smaller differences in variances can invalidate the F-test. Much more attention needs to be paid to unequal variances than to non-normality of data.