**Tram Le**

**Final:**

1. I use lm() function for fit linear regression model on education, then I find significant level for that model.

I notice the p value of Fertility, Agriculture, Catholic < .05. So these variables are significant

Examination and Infant.Mortality are not significant.

Then I use lm() again on just significant variables. After using dratfee to estimate education on only significant variables, I compute that the draftee is 11.78151. This means average a draftee can expect to have education is over 11 years. Also, after took all variables that not significant out of linear regression, t-value increase, higher F-stat and smaller p-value. So model look better.

* After using leveneTest I see varibles are not equal between 2 group. So we can't see ANOVA. So we will use oneway.test to see if spray affect on count or not
* After using oneway.test, there is significant level in count between group and still have no equality of variables, then I run Turkey's test to find specific differences.
* After using Tukey's test , there is a significant difference in some group but there is not a significant difference in other group. So run games\_howell\_testto find specific difference
* When I look at the specific differences between the groups using the Games Howell test, I see that there is a significant difference in some group but also there is not a significant difference in other group that was sprayed.
* Inclusion, result with all test, showing that type of spray have effect on count

1. After using logistic regression and knn algorithm to classify wine based on 13 variables. Both of these algorithms have an accuracy about 95%. This means that both model are able to well predict wine class with 95% samples in test\_wine. However the logictis regreesion model is slightly better. Because when I try changing the value of K for KNN model, this actually decreases accuracy a lot.
2. After remove wine class. From the cluster algorithm, the algorithm is able to identifies 3 different sets that means the algorithm notice there different group of wine in database.