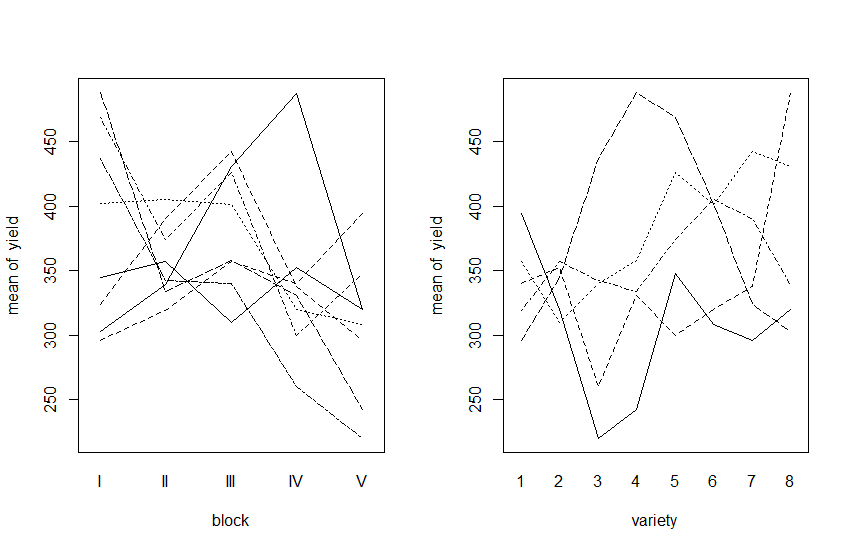
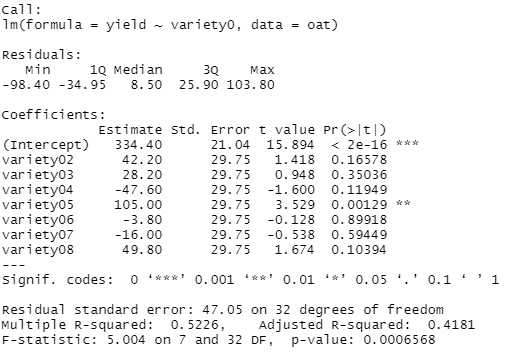
1.

(a) Two-way ANOVA

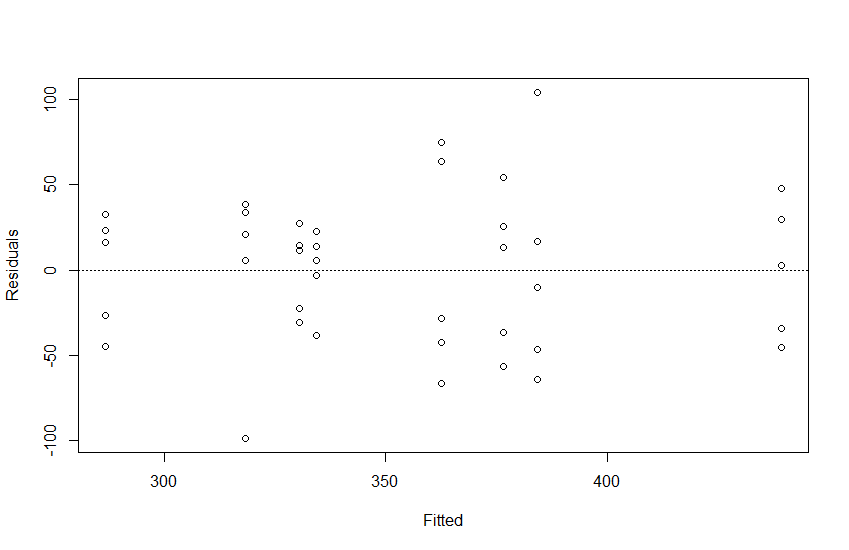
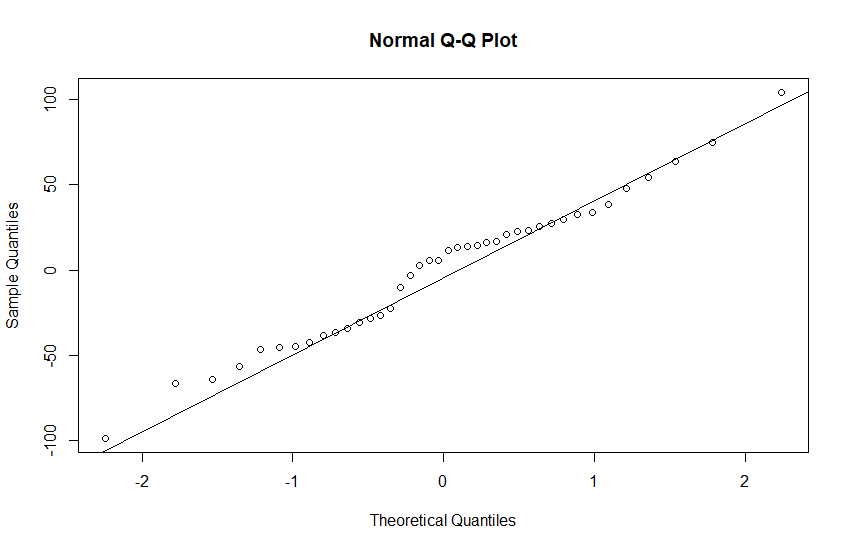
(b) In both cases the lines are not parallel, indicating interaction.

(c)



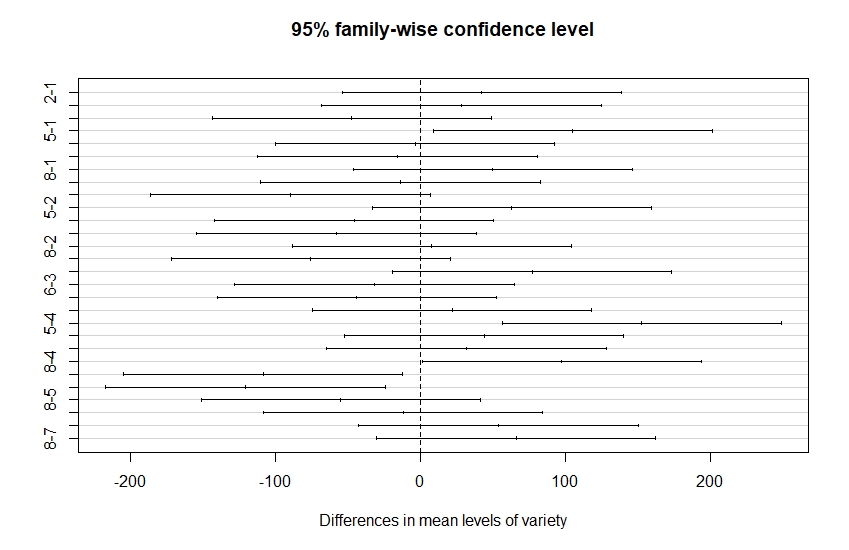
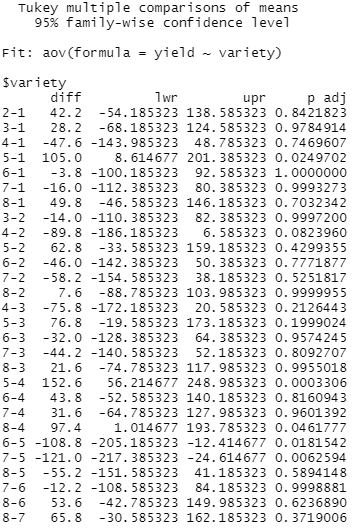
Use variety1 as the baseline, and we can see there is enough evidence to conclude that there is difference between variety1 and variety5. Therefore, the yield of oats is affected by different varieties at 5% significance level.

(d)



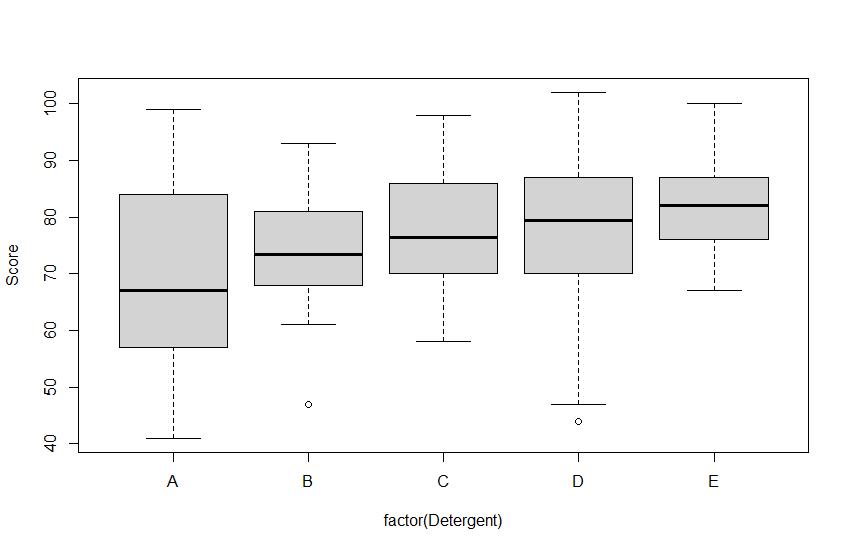
No, there isn’t any unusual finding.

(e)

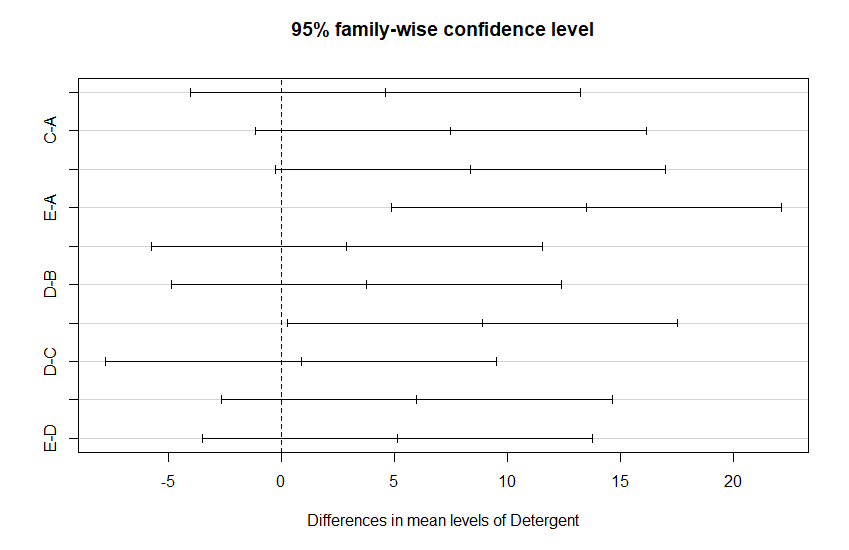


Yes, it’s necessary. There are differences between (variety1 and variety5), (variety4 and variety5), (variety4 and variety7), (variety4 and variety8), (variety5 and variety6), (variety5 and variety7).

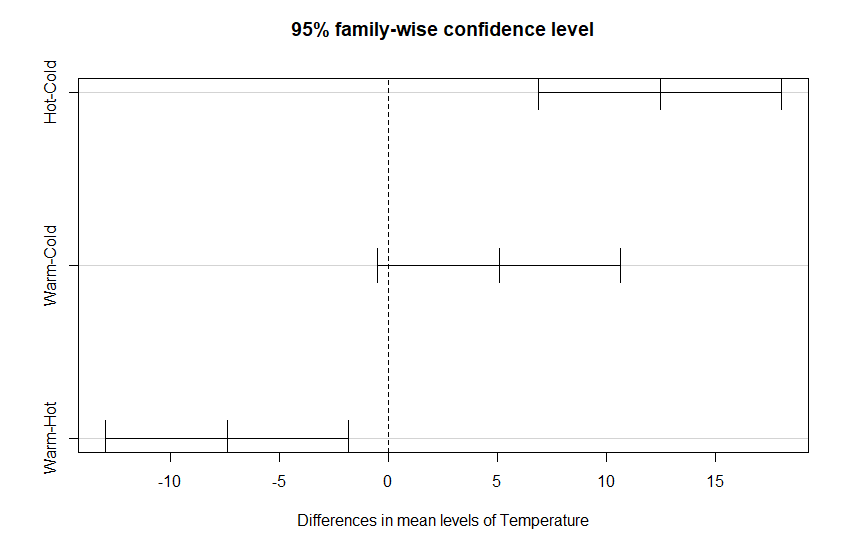
2.



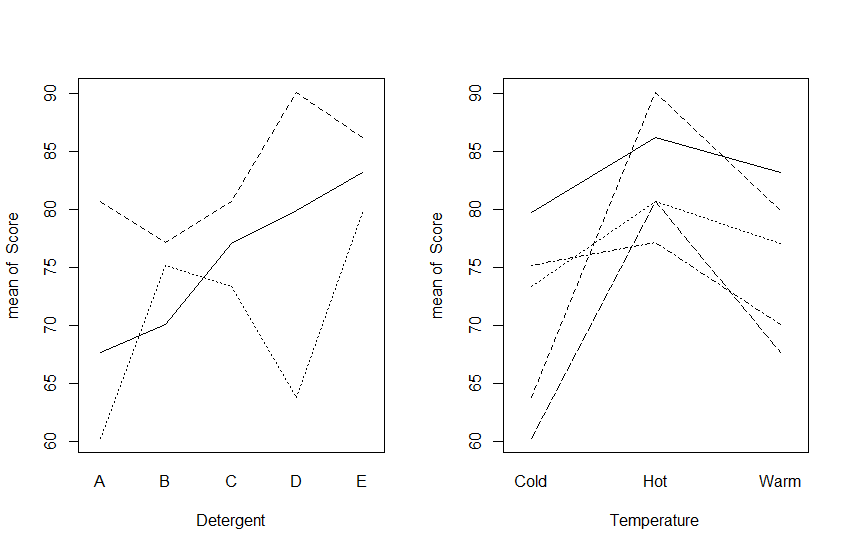
boxplot



There is sufficient statistical evidence to infer that there are differences in whiteness scores between (A and E) and (B and E) at 5% confidence level. However, There is no sufficient statistical evidence to infer that there are differences in whiteness scores between other detergents.



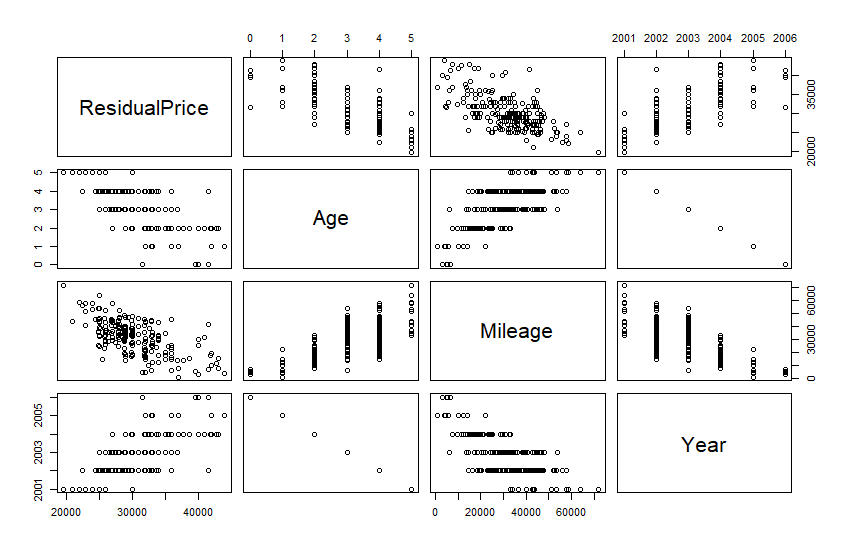
There is sufficient statistical evidence to infer that there are differences in whiteness scores between hot water and cold water.



In both cases the lines are not parallel, indicating interaction.

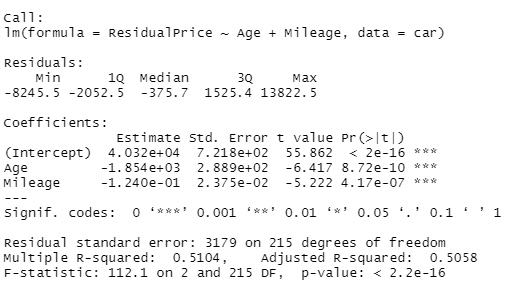
3.

(a)

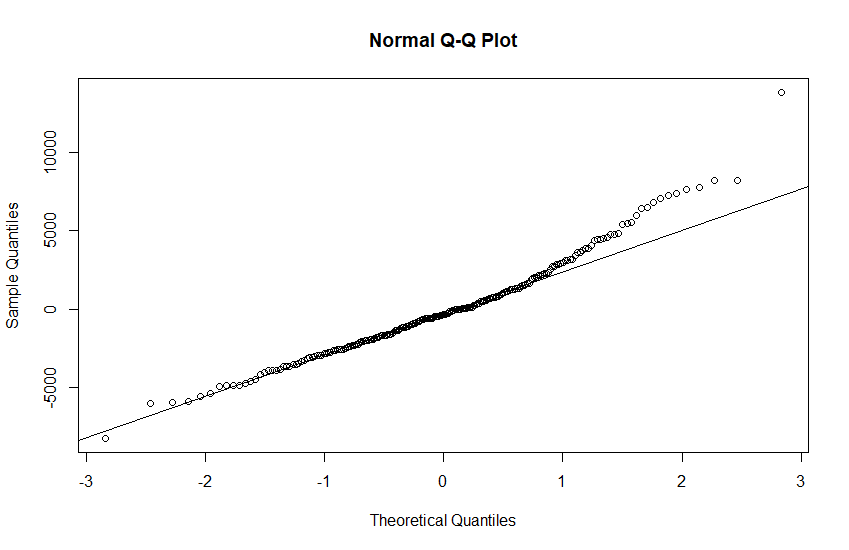
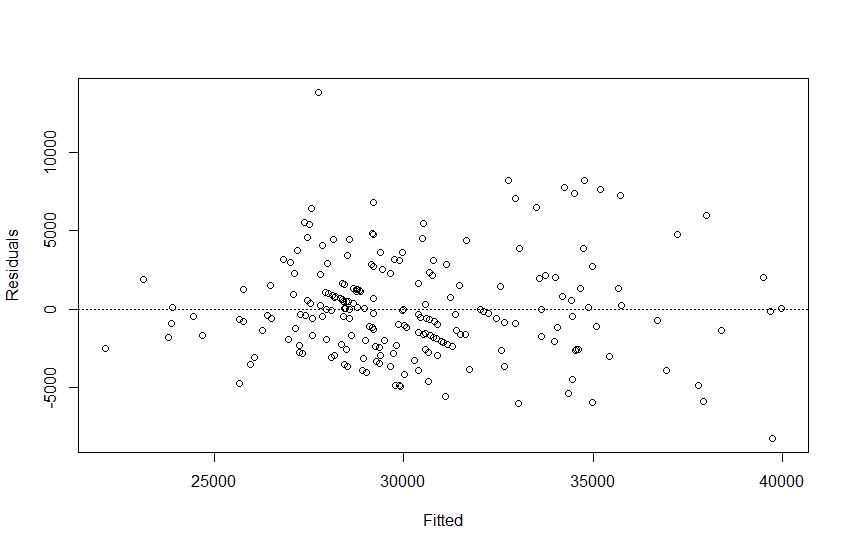


Yes, the relationships appear straight enough to permit using multiple linear regression with these variables. But the correlation between Age and Year is -1, so we need to drop one of them when we do multiple linear regression.

(b)



multiple linear regression model

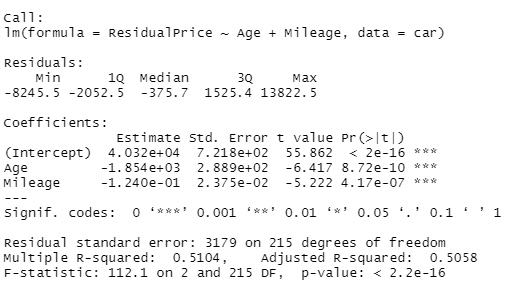


From normal Q-Q plot, we can conclude this model doesn’t meet the assumptions for multiple linear models.

(c)



(d)



The R-squared value:0.5104 means that there are about 51% of the variations can be explained by this model.

The p-value of Age is smaller than 0.05, which means we can reject the null hypothesis: coefficient of Age = 0. When Age increases one, the predicted price would decrease 1.854e+03.

The p-value of Mileage is smaller than 0.05, which means we can reject the null hypothesis: coefficient of Mileage = 0. When Mileage increases one, the predicted price would decrease 1.240e-01.

(e)

1. The brand and the quality of car

2. The habits of former driver

3. Whether there is an accident on it