Assignment 8 407

Ryan Klughart

09/04/2021

Question 1

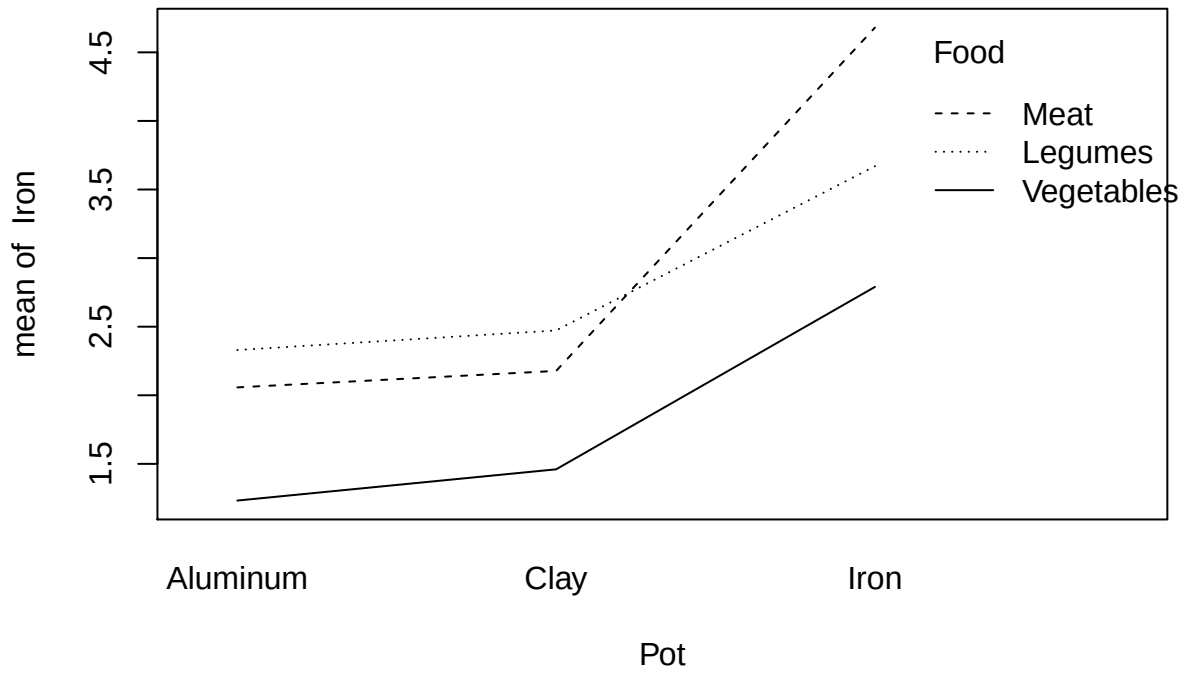
```
Food = c(rep("Meat",12),rep("Legumes",12),rep("Vegetables",12))
Pot = rep(c(rep("Aluminum",4),rep("Clay",4),rep("Iron",4)),3)
Iron = c(1.77,2.36,1.96,2.14,2.27,1.28,2.48,2.68,5.27,5.17,4.06,4.22,
         2.4,2.17,2.41,2.34,2.41,2.43,2.57,2.48,3.69,3.43,3.84,3.72,
         1.03,1.53,1.07,1.3,1.55,.79,1.68,1.82,2.45,2.99,2.8,2.92)


results = aov(Iron~Pot*Food)
anova(results)
```

```
## Analysis of Variance Table
##
## Response: Iron
##           Df Sum Sq Mean Sq F value  Pr(>F)
## Pot        2  24.89   12.45   92.26 8.5e-13 ***
## Food       2   9.30    4.65   34.46 3.7e-08 ***
## Pot:Food   4   2.64    0.66    4.89  0.0042 **
## Residuals 27   3.64    0.13
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The standard deviation of the errors is $\sqrt{MS_E} = \sqrt{.1349}$ =0.367
The main effect of food: The null hypothesis suggests that mean iron levels will not differ with respect to different foods. Since the F-value is large(92.26) with a p-value<0.05, we reject the null. This suggests that there is a significant difference in mean iron levels for at least two of the food types.
The main effect of pot: The null hypothesis suggests that mean iron levels will not differ with respect to different pot material. Since the F-value is large(34.46) with a p-value<0.05, we reject the null. This suggests that there is a significant difference in mean iron levels for at least two of the pot types.
The interaction of pot and food: The null hypothesis suggests that there is no interaction between food and pot. Since our F-value is large (4.83) and the p-value<0.05 we reject the null. This suggests that there is an interaction between food and pot. This interaction suggests that the effect of food type on iron levels will change from pot type to pot type, and vice-versa.

Question 2

```
interaction.plot(Pot,Food,Iron)
```




```
means = aggregate(Iron ~ Food*Pot, FUN=mean)
```

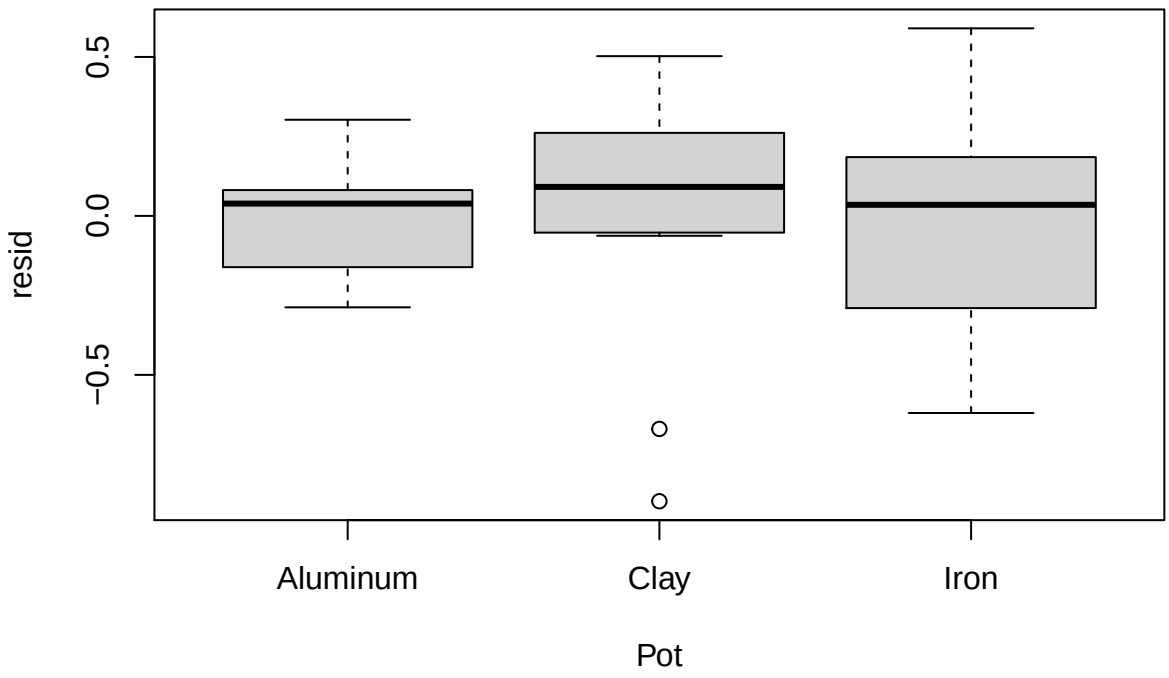
	Aluminum	Clay	Iron	Means
Meat	2.058	2.178	4.68	2.972
Legumes	2.33	2.473	3.67	2.824
Vegetables	1.232	1.46	2.79	1.827
Means	1.873	2.037	3.713	2.541

As we can see from the interaction plot, legumes and vegetables are parallel which suggests that the mean iron levels do not differ for vegetables and legumes when we change the pot material. However, meat does have a steeper positive slope than legumes and vegetables when we switch from a clay pot to an iron pot. This interaction tells us that iron levels will increase more so for meat than legumes and vegetables when changing the pot to an iron pot.

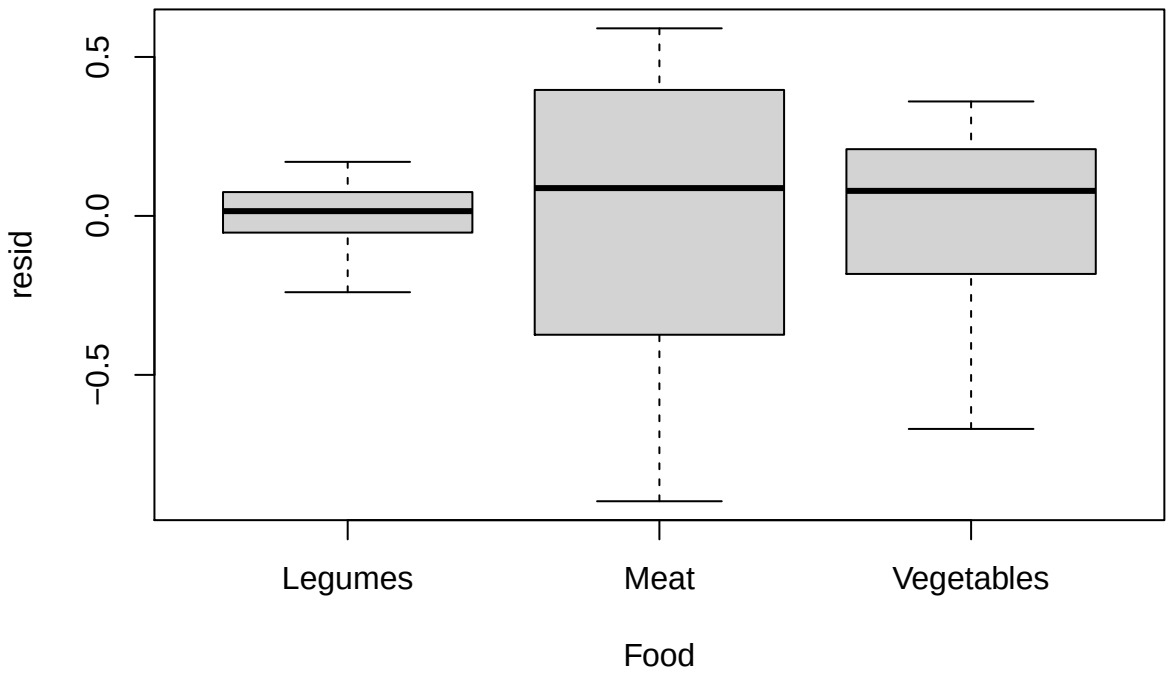
Question 3

a)

```
resid = residuals(results)

boxplot(resid ~ Pot)
```




```
boxplot(resid ~ Food)
```




```
sdFood = aggregate(resid ~ Food, FUN=sd)
sdFood
```

```
##         Food resid
## 1    Legumes 0.113
## 2       Meat 0.480
## 3 Vegetables 0.297
```

```
sdPot = aggregate(resid ~ Pot, FUN=sd)
sdPot
```

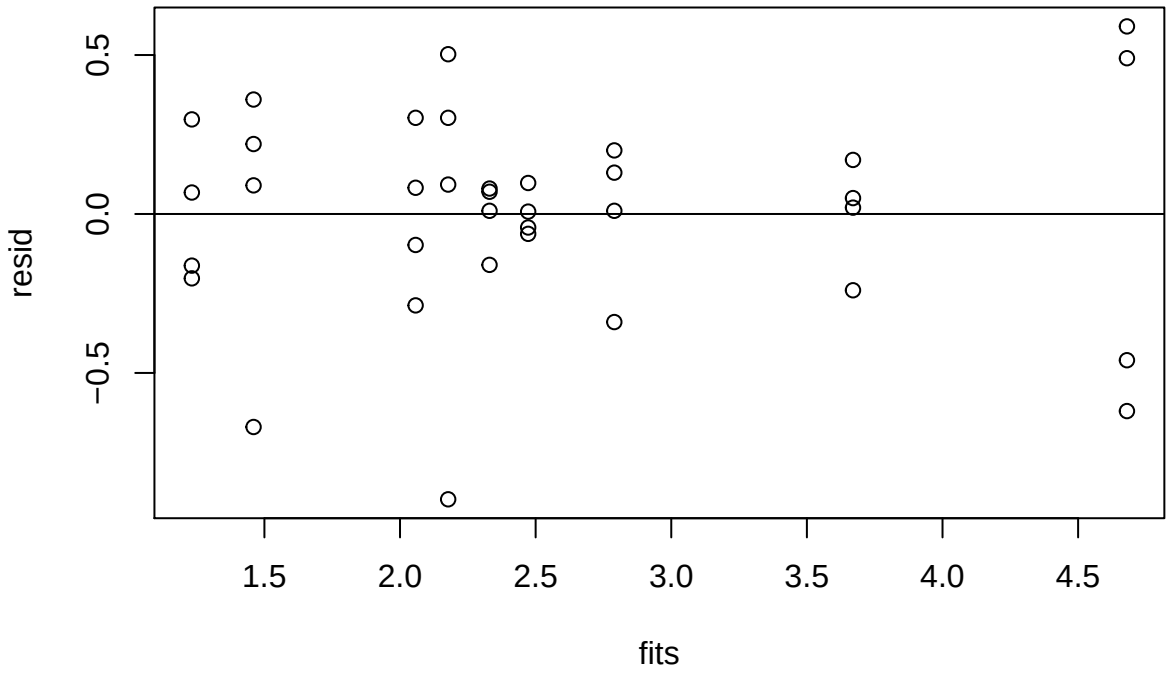
```
##        Pot resid
## 1 Aluminum 0.188
## 2     Clay 0.405
## 3     Iron 0.363
```

```
#boxplot(resid ~ Pot*Food,las=2)
```

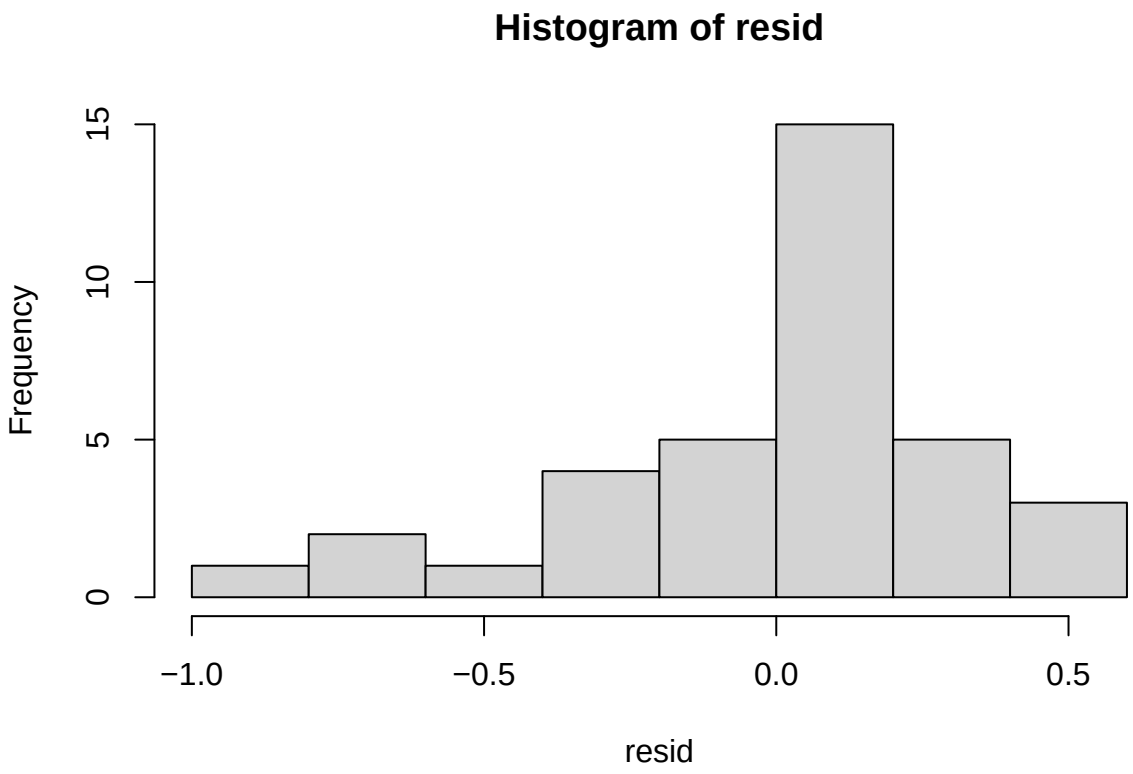
For the pot factor: As we can see from the side by side box plots, there is quite a lot of difference in variance between an iron pot, and the other two types of pots. There are also outliers on the low end of clay. Since the standard deviation of iron is greater than two times the standard deviation of Aluminum, the equal variance assumption is not satisfied.
For the food factor: As we can see from the box plots, legumes has a significantly smaller variance compared to meat and vegetables. Since the standard deviation of meat is greater than two times the standard deviation of legumes, the assumption of equal variance is not met.

b)

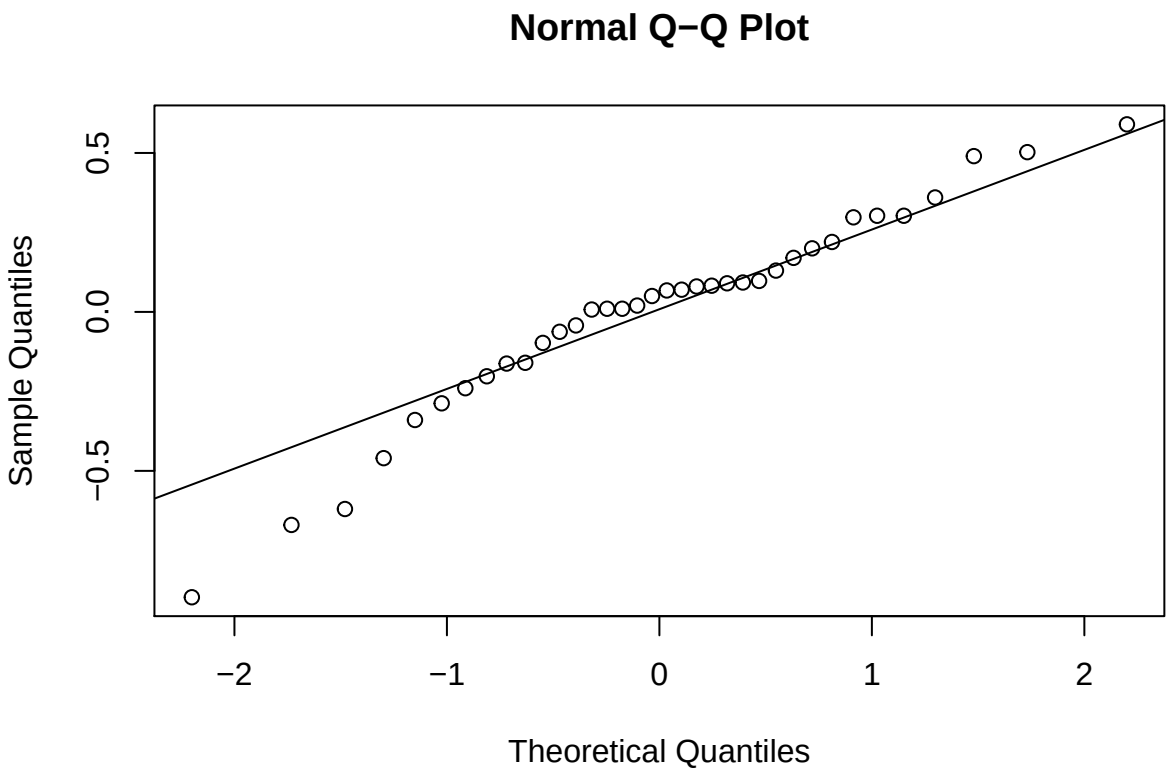
```
fits = fitted(results)
plot(resid~fits)
abline(a=0,b=0)
```




```
hist(resid)
```




```
qqnorm(y=resid)
qqline(y=resid)
```



From the histogram we can see that the residuals are normally distributed with a slight left skew. We can also see this in the Q-Q Plot as the lower tail deviates to the left of the line. From the residuals vs fitted plot we can see that there appears to be an outlier around 2.2, however this outlier does not appear serious as it is not too much further from the center than other values. Since the residuals appear to be normally distributed with no serious outliers the normality assumption is satisfied.