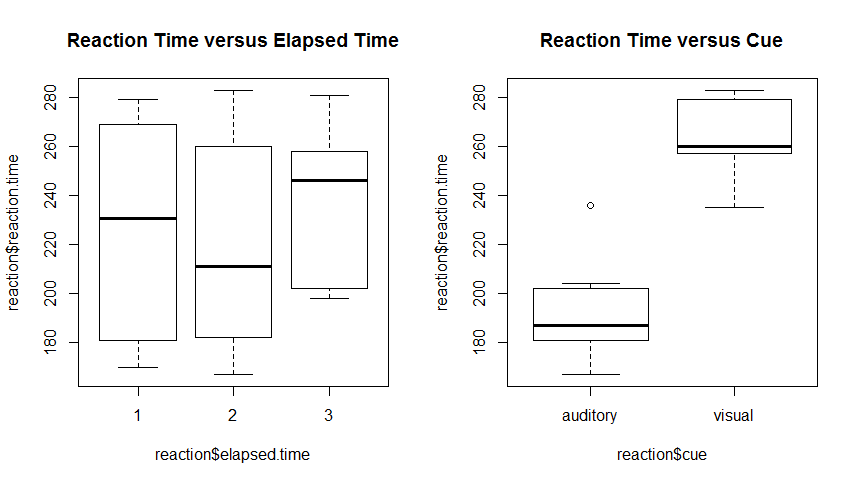
**Midterm 2**

Jiaqi Li

**1.1**

> plot(reaction.time~elapsed.time,main="Reaction Time versus Elapsed Time")

> plot(reaction.time~reaction$cue,main="Reaction Time versus Cue")



**1.2**

First, we want to set up the 2-way complete model:

Yijt = μ + αi + βj + (αβ)ij + ϵijt, ϵijt ∼ N(0,σ2) with iid

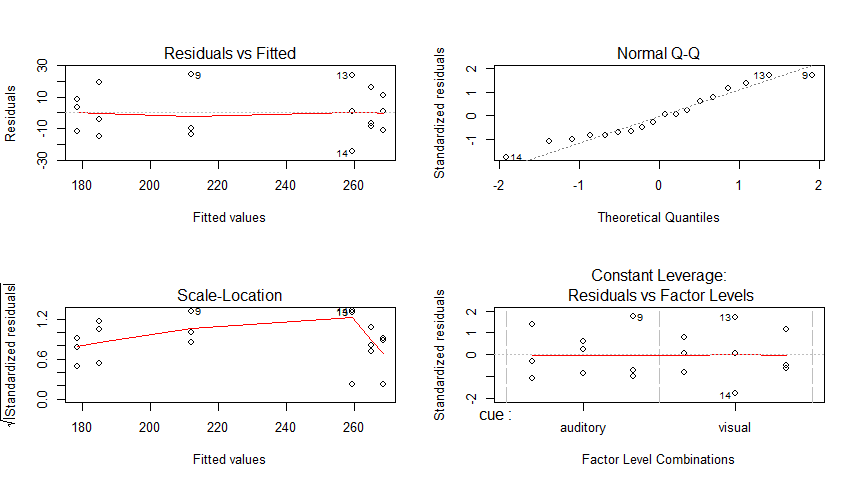
Yijt = the reaction time with ith cue and jth elapsed time.

i = auditory, visual j = 1, 2, 3 t = 1, …, ri

> lsm.reaction=aov(reaction.time~cue+elapsed.time+cue:elapsed.time,data=reaction)

> par(mfrow=c(2,2))

> plot(lsm.reaction)



The plots show that in the residuals plot of the 2-way complete model, constant variance is satisfied, and the Q-Q plot shows the normality. In this case, we do not need to do transformation. All assumptions are satisfied. We can do further study with this model.

|  |
| --- |
| > anova(lsm.reaction)  Analysis of Variance Table  Response: reaction.time  Df Sum Sq Mean Sq F value Pr(>F)  cue 1 23544.5 23544.5 81.3750 1.077e-06 \*\*\*  elapsed.time 2 1158.1 579.1 2.0013 0.1778  cue:elapsed.time 2 846.3 423.2 1.4626 0.2701  Residuals 12 3472.0 289.3  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 |

**For Interaction term**:

H0: combination of cue and elapsed time has no impact on the mean reaction time

Test Statistic F2,12 = 1.4626

p-value = 0.2701 > α = 0.05

Since the p-value of the interaction term is larger than 0.05, we fail to reject null hypothesis and conclude that different combinations of cue and elapsed time have no impact on mean reaction time.

**For Cue**:

H0: cue has no impact on reaction time

Test Statistic F1,12 = 81.3750

p-value = 1.077 × 10-6 < α = 0.05

Since the p-value of the cue is smaller than 0.05, we reject null hypothesis and conclude that different cues have impacts on mean reaction time.

|  |
| --- |
| > cld(reaction.cue,alpha=0.05)  cue lsmean SE df lower.CL upper.CL .group  auditory 191.8889 5.669934 12 179.5352 204.2426 1  visual 264.2222 5.669934 12 251.8685 276.5759 2  Results are averaged over the levels of: elapsed.time  Confidence level used: 0.95  significance level used: alpha = 0.05 |

By observing the chart and graph above, we can see that the visual cue has more impacts on the mean reaction time than the auditory cue has.

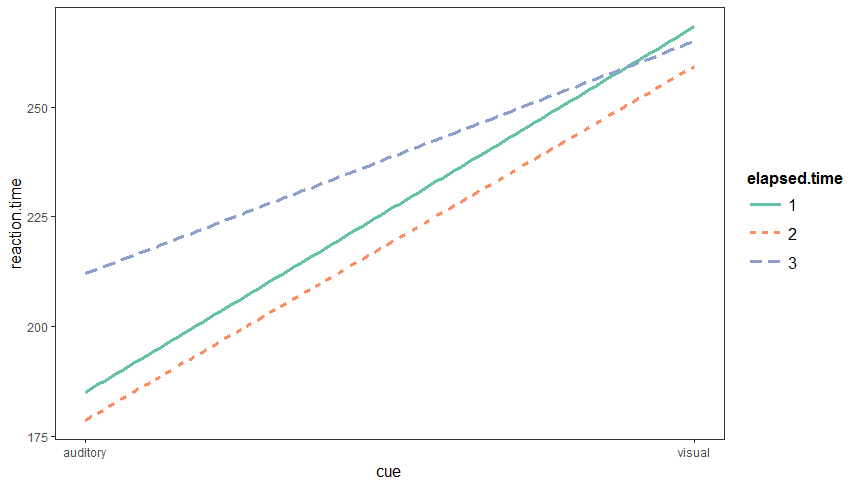
**For Elapsed Time**:

H0: elapsed time has no impact on reaction time

Test Statistic F1,14 = 1.3470

p-value = 0.2652 > α = 0.05

Since the p-value of the elapsed time is larger than 0.05, we fail to reject null hypothesis and conclude that different elapsed time have no impact on mean reaction time.



By observing the graph above, we can see that:

For auditory cue, elapsed time 1 and 2 have similar effect on reaction time but elapsed time 3 has a little bit more effects on reaction time than elapsed time 1 and 2 have.

For visual cue, elapsed time 1 and 3 have same effects on reaction time but elapsed time 2 has a little bit less effects on reaction time than elapsed time 1 and 3 have.

In general, combinations of cue and elapsed have no significant effects on reaction time.

For cue, auditory cue has less effects on reaction time than visual cue has.

For elapsed time, even though the elapsed time 2 has the smallest effects on reaction time, there is no significant different between elapsed times.

**2.**

Model: Yit = μ + τi + ϵit , ϵit ∼ N(0,σ2) with iid

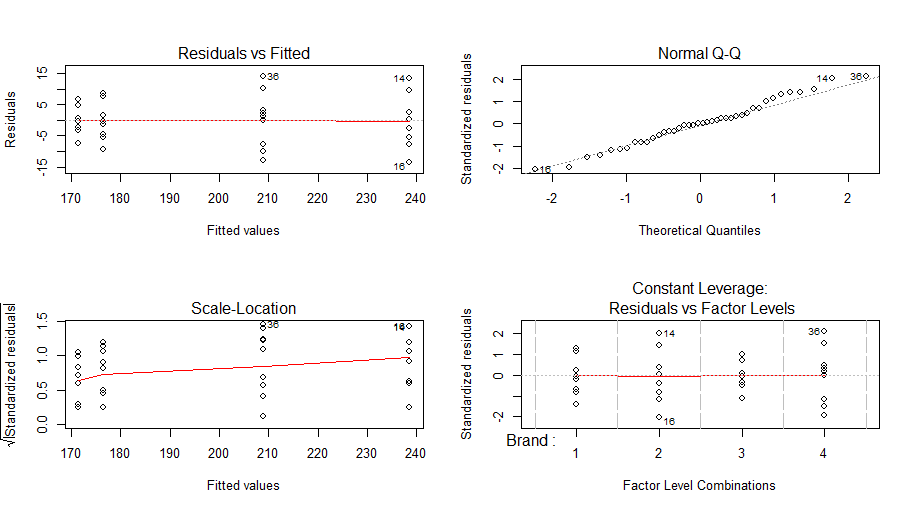
Yit = melting time of tth butter of ith brand.

i = 1, 2, 3, 4 t = 1, …, ri

> lsm.melt=aov(Times~Brand,data=melt)

> par(mfrow=c(2,2))

> plot(lsm.melt)



By observing the graph above, we can see that the residual plot shows a constant variance and the Q-Q plot shows the normality. Thus, all assumptions are satisfied. We can do further study.

|  |
| --- |
| > anova(lsm.melt)  Analysis of Variance Table  Response: Times  Df Sum Sq Mean Sq F value Pr(>F)  Brand 3 29385.7 9795.2 200.35 < 2.2e-16 \*\*\*  Residuals 36 1760.1 48.9  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 |

H0: Brand has no significant effect to the mean melting time.

H1: Brand has significant effects to the mean melting time.

Test-Statistic = F3,36 = 200.35

p-value < 2.2 × 10-16 < α =0.05

Since the p-value is smaller than 0.05, we reject null hypothesis and conclude that butter brands do have significant effect to the mean melting time.

**3.1**

This experiment is a completely randomized design because for each treatment (softwood and hardwood), 4 trees are randomly selected from a larger population of each species.

**3.2**

Model: Yijt = μ + αi + βj(i) + ϵijt , ϵijt ∼ N(0,σ2) with iid

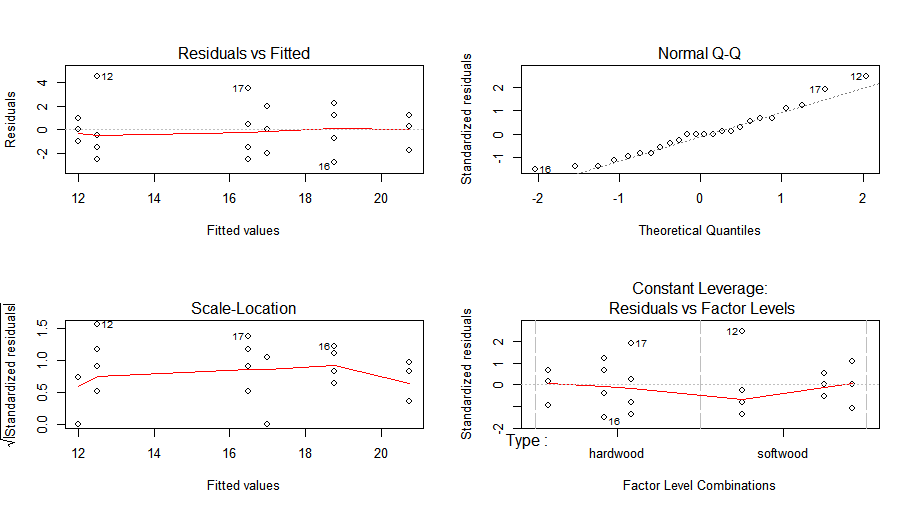
Yijt = nitrogen content of tth tree of ith type and jth species.

i = hardwood, softwood j = oak, ash, maple, pine, spruce, fir t = 1, … , ri

> lsm.wood=aov(Nconc~Type/Species,data=wood)

> par(mfrow=c(2,2))

> plot(lsm.wood)



The plots show that the residuals of the 2-way nested model satisfy constant variance and the Q-Q plot shows the normality. Thus, assumptions are satisfied. No transformation is needed at this time. We can do further study with this model.

|  |
| --- |
| > anova(lsm.wood)  Analysis of Variance Table  Response: Nconc  Df Sum Sq Mean Sq F value Pr(>F)  Type 1 140.167 140.167 31.7358 2.408e-05 \*\*\*  Type:Species 4 96.833 24.208 5.4811 0.004574 \*\*  Residuals 18 79.500 4.417  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 |

**For Interaction term**:

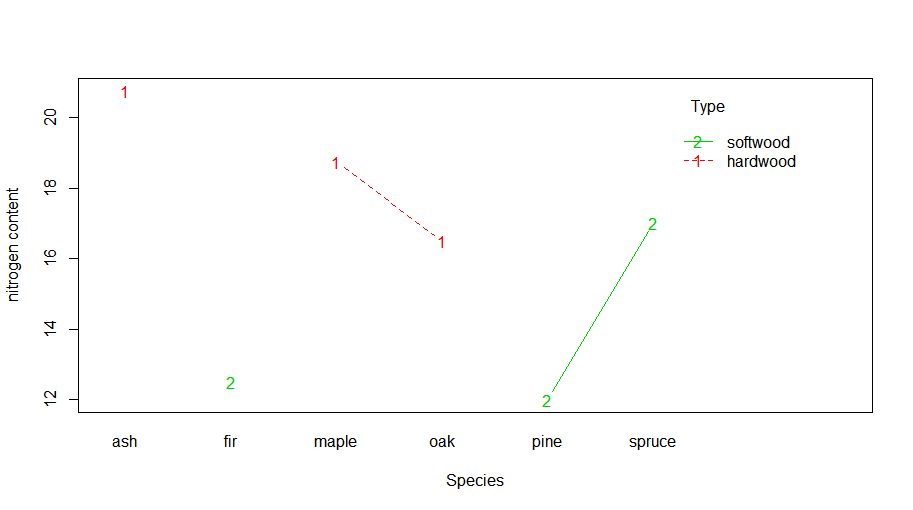
H0: combination of Type and Species has no impact on the mean reaction time

Test Statistic F4,18 = 5.4811

p-value = 0.004574 < α = 0.05

Since the p-value of the interaction term is smaller than 0.05, we reject null hypothesis and conclude that different combinations of Type and Species have impact on mean nitrogen content.

Then we will do pairwise comparison:



|  |
| --- |
| > contrast(wood.ST, method="pairwise")  contrast estimate SE df t.ratio p.value  ash,hardwood - maple,hardwood 2.00 1.486046 18 1.346 0.7567  ash,hardwood - oak,hardwood 4.25 1.486046 18 2.860 0.0923  ash,hardwood - fir,softwood 8.25 1.486046 18 5.552 0.0004  ash,hardwood - pine,softwood 8.75 1.486046 18 5.888 0.0002  ash,hardwood - spruce,softwood 3.75 1.486046 18 2.523 0.1689  maple,hardwood - oak,hardwood 2.25 1.486046 18 1.514 0.6602  maple,hardwood - fir,softwood 6.25 1.486046 18 4.206 0.0060  maple,hardwood - pine,softwood 6.75 1.486046 18 4.542 0.0029  maple,hardwood - spruce,softwood 1.75 1.486046 18 1.178 0.8416  oak,hardwood - fir,softwood 4.00 1.486046 18 2.692 0.1257  oak,hardwood - pine,softwood 4.50 1.486046 18 3.028 0.0670  oak,hardwood - spruce,softwood -0.50 1.486046 18 -0.336 0.9993  fir,softwood - pine,softwood 0.50 1.486046 18 0.336 0.9993  fir,softwood - spruce,softwood -4.50 1.486046 18 -3.028 0.0670  pine,softwood - spruce,softwood -5.00 1.486046 18 -3.365 0.0344  P value adjustment: tukey method for comparing a family of 6 estimates |

By observing the chart above, we can conclude that:

* Species of fir in soft wood category has less impact on mean nitrogen content than species of ash in hard wood category has.
* Species of pine in soft wood category has less impact on mean nitrogen content than species of ash in hard wood category has.
* Species of fir in soft wood category has less impact on mean nitrogen content than species of maple in hard wood category has.
* Species of pine in soft wood category has less impact on mean nitrogen content than species of maple in hard wood category has.
* Species of pine in soft wood category has more impact on mean nitrogen content than species of spruce in soft wood category has.
* No other comparisons are significantly different than zero based on the p-values.

**For Type**:

H0: Type has no impact on reaction time

Test Statistic F1,18 = 31.7358

p-value = 2.408 × 10-5 < α = 0.05

Since the p-value of the Type is smaller than 0.05, we reject null hypothesis and conclude that Type has impacts on mean nitrogen content.

|  |
| --- |
| > cld(wood.T, alpha=0.05)  Type lsmean SE df lower.CL upper.CL .group  softwood 13.83333 0.6066758 18 12.55875 15.10791 1  hardwood 18.66667 0.6066758 18 17.39209 19.94125 2  Results are averaged over the levels of: Species  Confidence level used: 0.95  significance level used: alpha = 0.05 |

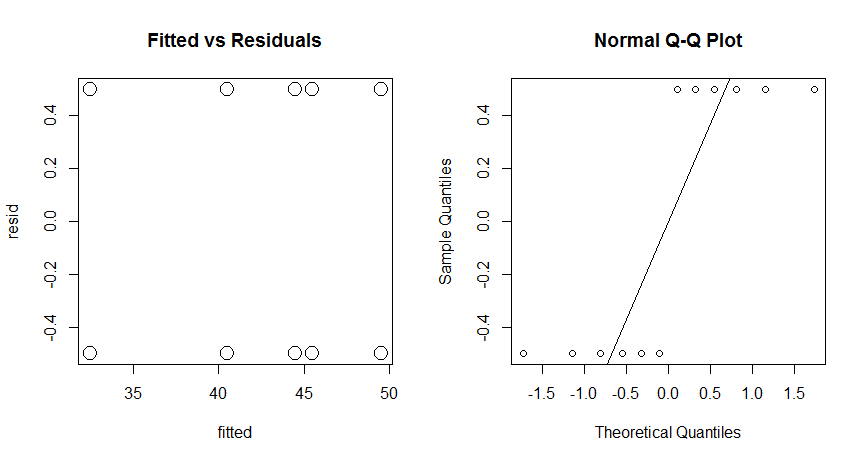
By observing the chart above, we can see that soft wood has less impacts on mean nitrogen content than hard wood has.

**4.1**

In this experiment, beans from each crocks are divided into 3 jars. This implies that each kind of beans in different crocks only have one jar. Then, it is not necessary to treat jar as a factor. Instead, we simply treat jar as a replicate.

(crocks nested in soak time, crocks crossed with recipes, response = reating, jars = replicates)

**4.2**



By observing the graph above, we can see that the constant variance is satisfied, but the Q-Q plot shows no normality. Thus, we want to try some transformations.

By trying transformation of 1/Rating, log(Rating), sqrt(Rating), the Q-Q plot gets better but the residual plots then show non-constant variance (transformations are not shown in the report but are included in the R code). Thus, we will still use the original model. Since not all assumptions are satisfied, we need to be careful with the p-value.

|  |
| --- |
| > Anova(lsm.taste,type="III")  Analysis of Deviance Table (Type III Wald chisquare tests)  Response: Rating  Chisq Df Pr(>Chisq)  (Intercept) 42672.667 1 < 2.2e-16 \*\*\*  SoakTim 450.667 1 < 2.2e-16 \*\*\*  Recipe 145.333 2 < 2.2e-16 \*\*\*  SoakTim:Recipe 81.333 2 < 2.2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 |

**For fixed Interaction between Soak Time and Recipes**:

H0: the interaction between Soak Time and Recipes has no effect on mean taste rating score

Test Statistic = χ22 = 81.333

p-value < 2.2 × 10-16 < α = 0.05

Since the p-value of the interaction between Soak Time and Recipes is smaller than 0.05, we reject null hypothesis and conclude that interaction between Soak Time and Recipes has significant effects on mean taste rating score. Also, since the interaction term has significant effects, we do not need to look at the main effect lines.

|  |
| --- |
| > difflsmeans(lsm.taste, "SoakTim:Recipe")  Differences of LSMEANS:  Estimate Standard Error DF t-value Lower CI Upper CI p-value  SoakTim:Recipe Long Barbecue - Short Barbecue 9 0.707 6 12.73 7.27 10.73 <2e-16 \*\*\*  SoakTim:Recipe Long Barbecue - Long Original 4 0.707 6 5.66 2.27 5.73 0.001 \*\*  SoakTim:Recipe Long Barbecue - Short Original 17 0.707 6 24.04 15.27 18.73 <2e-16 \*\*\*  SoakTim:Recipe Long Barbecue - Long Refried 5 0.707 6 7.07 3.27 6.73 4e-04 \*\*\*  SoakTim:Recipe Long Barbecue - Short Refried 9 0.707 6 12.73 7.27 10.73 <2e-16 \*\*\*  SoakTim:Recipe Short Barbecue - Long Original -5 0.707 6 -7.07 -6.73 -3.27 4e-04 \*\*\*  SoakTim:Recipe Short Barbecue - Short Original 8 0.707 6 11.31 6.27 9.73 <2e-16 \*\*\*  SoakTim:Recipe Short Barbecue - Long Refried -4 0.707 6 -5.66 -5.73 -2.27 0.001 \*\*  SoakTim:Recipe Short Barbecue - Short Refried 0 0.707 6 0.00 -1.73 1.73 1.000  SoakTim:Recipe Long Original - Short Original 13 0.707 6 18.38 11.27 14.73 <2e-16 \*\*\*  SoakTim:Recipe Long Original - Long Refried 1 0.707 6 1.41 -0.73 2.73 0.207  SoakTim:Recipe Long Original - Short Refried 5 0.707 6 7.07 3.27 6.73 4e-04 \*\*\*  SoakTim:Recipe Short Original - Long Refried -12 0.707 6 -16.97 -13.73 -10.27 <2e-16 \*\*\*  SoakTim:Recipe Short Original - Short Refried -8 0.707 6 -11.31 -9.73 -6.27 <2e-16 \*\*\*  SoakTim:Recipe Long Refried - Short Refried 4 0.707 6 5.66 2.27 5.73 0.001 \*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1 |

By observing the chart above, we can conclude that:

* Beans with Short soaking time and Barbecue recipes have less impacts on the mean taste rating score than beans with Long soaking time and Barbecue recipes.
* Beans with Long soaking time and Original recipes have less impacts on the mean taste rating score than beans with Long soaking time and Barbecue recipes.
* Beans with Short soaking time and Original recipes have less impacts on the mean taste rating score than beans with Long soaking time and Barbecue recipes.
* Beans with Long soaking time and Refried recipes have less impacts on the mean taste rating score than beans with Long soaking time and Barbecue recipes.
* Beans with Short soaking time and Refried recipes have less impacts on the mean taste rating score than beans with Long soaking time and Barbecue recipes.
* Beans with Short soaking time and Barbecue recipes have less impacts on the mean taste rating score than beans with Long soaking time and Original recipes.
* Beans with Short soaking time and Original recipes have less impacts on the mean taste rating score than beans with Short soaking time and Barbecue recipes.
* Beans with Short soaking time and Barbecue recipes have less impacts on the mean taste rating score than beans with Long soaking time and Refried recipes.
* Beans with Short soaking time and Original recipes have less impacts on the mean taste rating score than beans with Long soaking time and Original recipes.
* Beans with Short soaking time and Refried recipes have less impacts on the mean taste rating score than beans with Long soaking time and Original recipes.
* Beans with Short soaking time and Original recipes have less impacts on the mean taste rating score than beans with Long soaking time and Refried recipes.
* Beans with Short soaking time and Original recipes have less impacts on the mean taste rating score than beans with Short soaking time and Refried recipes.
* Beans with Short soaking time and Refried recipes have less impacts on the mean taste rating score than beans with Long soaking time and Refried recipes.
* No other comparisons are significantly different than zero based on the p-values.

**For random Interaction between Soak Time and Crock**:

|  |
| --- |
| > rand(lsm.taste)  Analysis of Random effects Table:  Chi.sq Chi.DF p.value  Crock:SoakTim -8.55 1 1 |

H0: σ2crock = 0

Test Statistic = χ21 = -8.55

p-value = 1 > α = 0.05

Since the p-value of the interaction between Soak Time and Crock is larger than 0.05, we fail to reject null hypothesis and conclude that there are no differences between Crock.

**R code:**

**#install.packages("lmerTest")**

**#install.packages("lsmeans")**

**#install.packages("car")**

**#install.packages("multcompView")**

**#install.packages("lme4")**

**#install.packages("jtools")**

**library(lsmeans)**

**library(car)**

**library(multcompView)**

**library(lme4)**

**library(lmerTest)**

**library(jtools)**

**options(contrasts =c("contr.sum", "contr.poly"))**

**#1.data-------------------------------------------------------------------------------------------**

**cue=c(rep("auditory",9),rep("visual",9))**

**elapsed.time=rep(rep(1:3,each=3),2)**

**reaction.time=c(204,170,181,167,182,187,202,198,236,257,279,269,283,235,260,256,281,258)**

**#1.1-------------------------------------------------------------------------------------------**

**reaction=data.frame(cue=cue,elapsed.time=as.factor(elapsed.time),reaction.time=reaction.time)**

**par(mfrow=c(1,2))**

**plot(reaction$reaction.time~reaction$elapsed.time,main="Reaction Time versus Elapsed Time")**

**plot(reaction$reaction.time~reaction$cue,main="Reaction Time versus Cue")**

**#1.2-------------------------------------------------------------------------------------------**

**lsm.reaction=aov(reaction.time~cue+elapsed.time+cue:elapsed.time,data=reaction)**

**par(mfrow=c(2,2))**

**plot(lsm.reaction)**

**anova(lsm.reaction)**

**interact\_plot(lsm.reaction,pred="cue",modx="elapsed.time")**

**unloadNamespace("lmerTest")**

**library(lsmeans)**

**reaction.cue=lsmeans(lsm.reaction,~cue)**

**cld(reaction.cue,alpha=0.05)**

**#2-------------------------------------------------------------------------------------------**

**Brand<-c(rep(1,10), rep(2, 10), rep(3,10), rep(4, 10))**

**Times<-c(167, 171, 178, 175, 184, 176, 185, 172, 178, 178,**

**231, 233, 236, 252, 233, 225, 241, 248, 239, 248,**

**176, 168, 171, 172, 178, 176, 169, 164, 169, 171,**

**201, 199, 196, 211, 209, 223, 209, 219, 212, 210)**

**melt=data.frame(Brand=as.factor(Brand),Times=Times)**

**lsm.melt=aov(Times~Brand,data=melt)**

**summary(lsm.melt)**

**par(mfrow=c(2,2))**

**plot(lsm.melt)**

**anova(lsm.melt)**

**#3.2-------------------------------------------------------------------------------------------**

**wood=read.table("wood.csv",header=TRUE)**

**wood**

**lsm.wood=aov(Nconc~Type/Species,data=wood)**

**par(mfrow=c(2,2))**

**plot(lsm.wood)**

**anova(lsm.wood)**

**par(mfrow=c(1,1))**

**interaction.plot(x.factor = wood$Species, trace.factor = wood$Type, response = wood$Nconc,**

**type ="b",col = 2:3,xlab ="Species", ylab ="nitrogen content", trace.label ="Type")**

**wood.ST=lsmeans(lsm.wood, ~Species:Type)**

**contrast(wood.ST, method="pairwise")**

**wood.T=lsmeans(lsm.wood, ~Type)**

**cld(wood.T, alpha=0.05)**

**#4.2--------------------------------------------------------------------------------------------**

**beans=read.table("Beans.csv",header=TRUE)**

**beans**

**Crock=as.factor(beans$Crock)**

**SoakTim=beans$SoakTim**

**Recipe=beans$Recipe**

**Jar=as.factor(beans$Jar)**

**Rating=beans$Rating**

**beans.frame=data.frame(Crock=Crock,SoakTim=SoakTim,Recipe=Recipe,Jar=Jar,Rating=Rating)**

**library(lmerTest)**

**lsm.taste=lmer(Rating~SoakTim+(1|Crock:SoakTim)+Recipe+SoakTim:Recipe,data=beans.frame)**

**par(mfrow=c(1,2))**

**fitted=fitted(lsm.taste)**

**resid=residuals(lsm.taste)**

**plot(fitted,resid,main="Fitted vs Residuals",pch=1,cex=2)**

**qqnorm(resid)**

**qqline(resid)**

**lsm.taste1=lmer(1/(Rating)~SoakTim+(1|Crock:SoakTim)+Recipe+SoakTim:Recipe,data=beans.frame)**

**fitted1=fitted(lsm.taste1)**

**resid1=residuals(lsm.taste1)**

**plot(fitted1,resid1,main="Fitted vs Residuals",pch=1,cex=2)**

**qqnorm(resid1)**

**qqline(resid1)**

**lsm.taste2=lmer(sqrt(Rating)~SoakTim+(1|Crock:SoakTim)+Recipe+SoakTim:Recipe,data=beans.frame)**

**fitted2=fitted(lsm.taste2)**

**resid2=residuals(lsm.taste2)**

**plot(fitted2,resid2,main="Fitted vs Residuals",pch=1,cex=2)**

**qqnorm(resid2)**

**qqline(resid2)**

**lsm.taste3=lmer(log(Rating)~SoakTim+(1|Crock:SoakTim)+Recipe+SoakTim:Recipe,data=beans.frame)**

**fitted3=fitted(lsm.taste3)**

**resid3=residuals(lsm.taste3)**

**plot(fitted3,resid3,main="Fitted vs Residuals",pch=1,cex=2)**

**qqnorm(resid3)**

**qqline(resid3)**

**Anova(lsm.taste,type="III")**

**difflsmeans(lsm.taste, "SoakTim:Recipe")**

**rand(lsm.taste)**