

Lista 3 - MAE0327

Guilherme N^oUSP: 8943160 e Leonardo N^oUSP: 9793436

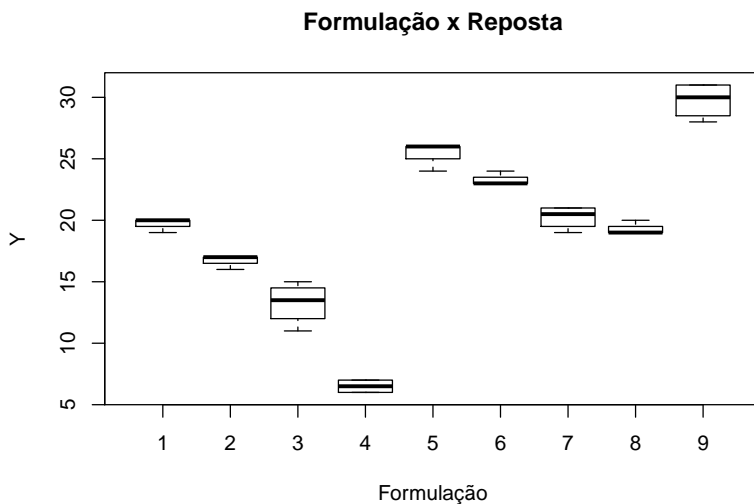
Exercício 1

Um engenheiro químico deseja avaliar a eficácia de nove formulações alternativas de um detergente em termos da duração da espuma durante o uso. Três pias estavam disponíveis e três pessoas foram instruídas a usar as pias para lavar pratos a uma taxa constante. Cada bloco consistia em três unidades experimentais em que a unidade experimental foi uma das pias com uma quantidade fixa de água potável e uma quantidade fixa de sujeira agregada. Três formulações de detergentes foram aleatoriamente designadas para as três pias em cada bloco. A variável resposta (Y) foi a duração da espuma medida pelo número de pratos lavados antes dela desaparecer. O engenheiro utilizou o desenho de blocos incompletos balanceados que segue. Analise os dados deste experimento e tire conclusões.

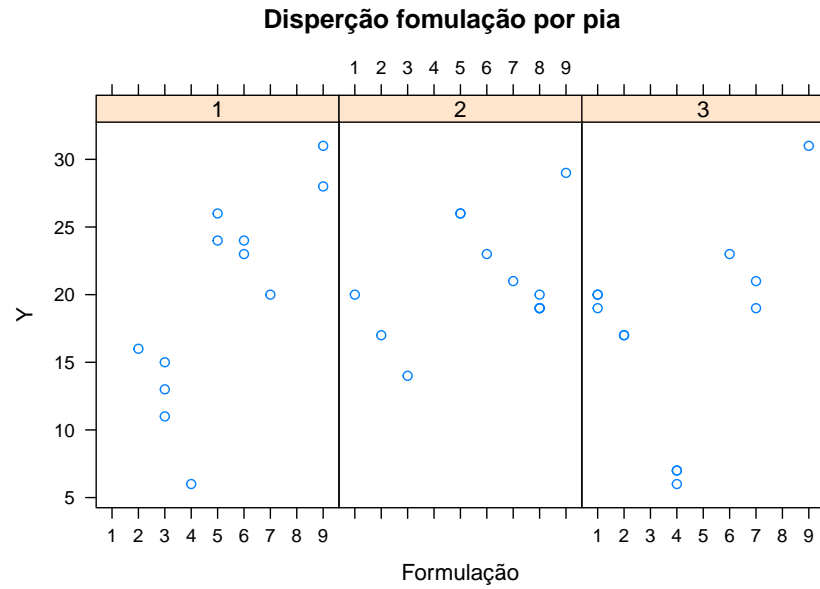
Bloco	Formulação			Y		
	Pia 1	Pia 2	Pia 3	Pia 1	Pia 2	Pia 3
1	3	8	4	13	20	7
2	4	9	2	6	29	17
3	3	6	9	15	23	31
4	9	5	1	31	26	20
5	2	7	6	16	21	23
6	6	5	4	23	26	6
7	9	8	7	28	19	21
8	7	1	4	20	20	7
9	6	8	1	24	19	20
10	5	8	2	26	19	17
11	5	3	7	24	14	19
12	3	2	1	11	17	19

Resolução

Primeiramente iremos fazer uma análise descritiva, assim:



Com o boxplot acima podemos notar que a variabilidade das formulações estão relativamente controladas e que os valores medianos são bem diferentes entre si.



E com o gráfico acima é possível confirmar que possuem valores diferentes por formulação em cada pia separadamente.

E a seguinte tabela reforça ainda essas conclusões:

Formulação	Média	Desvio Padrão	Mediana
1	19.75	0.5	20
2	16.75	0.5	17
3	13.25	1.708	13.5
4	6.5	0.577	6.5
5	25.5	1	26
6	23.25	0.5	23
7	20.25	0.957	20.5
8	19.25	0.5	19
9	29.75	1.5	30

Feito a análise descritiva, iremos fazer o ajuste do modelo onde temos que:

$K = 3$ (número de unidades experimentais por bloco)

$t = 9$ (número de tratamentos)

$b = 12$ (número de blocos)

$r = 3$ (número de repetições do tratamento)

$\lambda = 1$ (número de vezes em que 2 tratamentos aparecem juntos num mesmo bloco)

$N = 36$ (número total de unidades experimentais)

Obtemos a tabela de ANOVA:

Table 1: Tabela de ANOVA

	Sum Sq	Df	F value	Pr(>F)
(Intercept)	347.175	1	382.984	0.000
Bloco	10.783	11	1.081	0.437
Formulacao	993.753	8	137.031	0.000
Pia	1.828	2	1.008	0.390
Residuals	12.691	14	NA	NA

onde pode-se ver que a formulação é um efeito significativo para o experimento. E o Bloco e a Pia podem entrar como fatores de controle, como é feito a seguir:

```
##
## Error: Bloco
##           Df Sum Sq Mean Sq F value Pr(>F)
## Formulacao  8     408    51.00   30.77 0.00848 **
## Residuals   3         5     1.66
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Error: Bloco:Pia
##           Df Sum Sq Mean Sq F value Pr(>F)
## Formulacao  8 1082.8   135.35   149.2 1.5e-13 ***
## Residuals  16    14.5     0.91
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Nota-se que se considerarmos o Bloco e a Pia como fator de erro, a Formulação é mais significativa se considerarmos apenas o Bloco. Continuando a análise obtemos:

```
##
## ANALYSIS BIB: Y
## Class level information
##
## Block:  1 2 3 4 5 6 7 8 9 10 11 12
## Trt  :  3 8 4 9 2 6 5 1 7
##
## Number of observations:  36
##
## Analysis of Variance Table
##
## Response: Y
##           Df Sum Sq Mean Sq F value Pr(>F)
## block.unadj 11   412.97   37.543   41.374 1.250e-09 ***
## trt.adj      8 1082.81  135.352  149.163 1.495e-13 ***
## Residuals   16    14.52    0.907
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## coefficient of variation: 4.9 %
```

```

## Y Means: 19.36111
##
## Formulacao, statistics
##
##      Y mean.adj      SE r      std Min Max
## 1 19.75 19.694444 0.5422796 4 0.5000000 19 20
## 2 16.75 17.138889 0.5422796 4 0.5000000 16 17
## 3 13.25 13.361111 0.5422796 4 1.7078251 11 15
## 4  6.50  6.472222 0.5422796 4 0.5773503  6  7
## 5 25.50 25.472222 0.5422796 4 1.0000000 24 26
## 6 23.25 22.916667 0.5422796 4 0.5000000 23 24
## 7 20.25 20.583333 0.5422796 4 0.9574271 19 21
## 8 19.25 19.138889 0.5422796 4 0.5000000 19 20
## 9 29.75 29.472222 0.5422796 4 1.5000000 28 31
##
## Tukey
## Alpha      : 0.05
## Std.err    : 0.5499719
## HSD        : 2.766913
## Parameters BIB
## Lambda     : 1
## treatmeans : 9
## Block size : 3
## Blocks     : 12
## Replication: 4
##
## Efficiency factor 0.75
##
## <<< Book >>>
##
## Comparison between treatments means
##      Difference pvalue sig.
## 1 - 2  2.5555556 0.0823  .
## 1 - 3  6.3333333 0.0000 ***
## 1 - 4 13.2222222 0.0000 ***
## 1 - 5 -5.7777778 0.0000 ***
## 1 - 6 -3.2222222 0.0164  *
## 1 - 7 -0.8888889 0.9578
## 1 - 8  0.5555556 0.9978
## 1 - 9 -9.7777778 0.0000 ***
## 2 - 3  3.7777778 0.0041  **
## 2 - 4 10.6666667 0.0000 ***
## 2 - 5 -8.3333333 0.0000 ***
## 2 - 6 -5.7777778 0.0000 ***
## 2 - 7 -3.4444444 0.0095  **
## 2 - 8 -2.0000000 0.2703
## 2 - 9 -12.3333333 0.0000 ***
## 3 - 4  6.8888889 0.0000 ***
## 3 - 5 -12.1111111 0.0000 ***
## 3 - 6 -9.5555556 0.0000 ***
## 3 - 7 -7.2222222 0.0000 ***
## 3 - 8 -5.7777778 0.0000 ***
## 3 - 9 -16.1111111 0.0000 ***
## 4 - 5 -19.0000000 0.0000 ***

```

```

## 4 - 6 -16.4444444 0.0000 ***
## 4 - 7 -14.1111111 0.0000 ***
## 4 - 8 -12.6666667 0.0000 ***
## 4 - 9 -23.0000000 0.0000 ***
## 5 - 6 2.5555556 0.0823 .
## 5 - 7 4.8888889 0.0003 ***
## 5 - 8 6.3333333 0.0000 ***
## 5 - 9 -4.0000000 0.0024 **
## 6 - 7 2.3333333 0.1358
## 6 - 8 3.7777778 0.0041 **
## 6 - 9 -6.5555556 0.0000 ***
## 7 - 8 1.4444444 0.6484
## 7 - 9 -8.8888889 0.0000 ***
## 8 - 9 -10.3333333 0.0000 ***
##
## Treatments with the same letter are not significantly different.
##
##          Y groups
## 9 29.472222      a
## 5 25.472222      b
## 6 22.916667     bc
## 7 20.583333     cd
## 1 19.694444     de
## 8 19.138889     de
## 2 17.138889      e
## 3 13.361111      f
## 4 6.472222       g

```

Pelas comparações múltiplas de Tukey, agrupamos as formulações em 4 grupos:

A nona formulação que possui valores mais altos, em seguida dos valores da quinta, sexta, sétima, primeira, oitava e segunda formulações que são maiores que a terceira formulação, que por sua vez é maior que a quarta formulação.

Exercício 2

Explique como a relação $\lambda(t-1) = r(k-1)$ foi definida num planejamento em blocos incompletos balanceados. Consultar Kuehl, R. O. (1999). *Design of Experiments: Statistical Principles of Research Design and Analysis*. 2nd ed. Duxbury Press.

Resolução

A relação $\lambda(t-1) = r(k-1)$ foi definida num planejamento em blocos incompletos balanceados pelo fato de que todos os tratamentos sejam igualmente replicados e cada par de tratamento no mesmo bloco o que resulta em igual precisão para todas as comparações. Assim o delineamento que possui r repetições de tratamentos t em b blocos de k unidades experimentais com $k < t$, terá o número total de unidades experimentais de $N = rt = bk$. O número de blocos em que cada par de tratamentos ocorre juntos é $\lambda = r(k-1)/(t-1)$ em que $\lambda < r < b$. O valor de λ vem do fato de que cada tratamento está pareado com outros $t-1$ tratamentos no delineamento λ vezes e existem $\lambda(t-1)$ pares pra algum tratamento no experimento, assim o mesmo tratamento aparece em r blocos com $k-1$ outros tratamentos, e cada tratamento aparece em $r(k-1)$ pares. Então

$$\lambda(t-1) = r(k-1) \Rightarrow \lambda = \frac{r(k-1)}{t-1}$$

Exercício 3

Um experimento foi conduzido para determinar o efeito da temperatura e do tempo de aquecimento na força de tração do aço. Duas temperaturas e três tempos de aquecimento foram selecionados. O experimento foi conduzido da seguinte forma: um forno foi aquecido a uma temperatura escolhida ao acaso e três porções de aço foram colocadas no forno. Depois de 10 minutos uma porção foi removida, depois de 20 minutos outra porção foi removida, e depois de 30 minutos a última porção foi removida. Em seguida, a temperatura do forno foi mudada para o outro nível e o processo foi repetido. Todas as porções foram então submetidas à tração. Para coletar todos os dados, exibidos na tabela a seguir, foi necessário repetir o experimento 4 vezes.

Tabela. Dados			
Repetição	Tempo (minutos)	Temperatura (graus F)	
		1500	1600
1	10	63	89
	20	54	91
	30	61	62
2	10	50	80
	20	52	72
	30	59	69
3	10	48	73
	20	74	81
	30	71	69
4	10	54	88
	20	48	92
	30	59	64

- (a) Qual planejamento você pode associar a este experimento? Defina a unidade experimental.

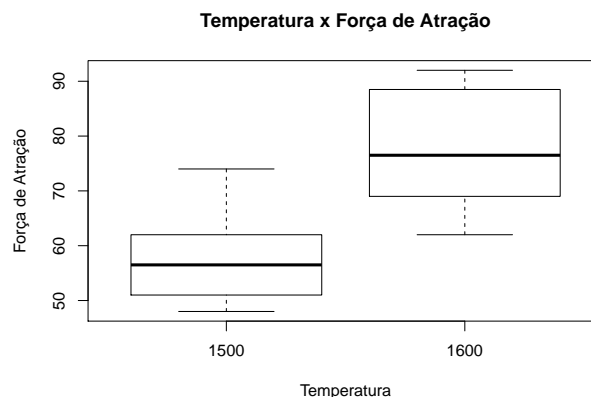
Resolução

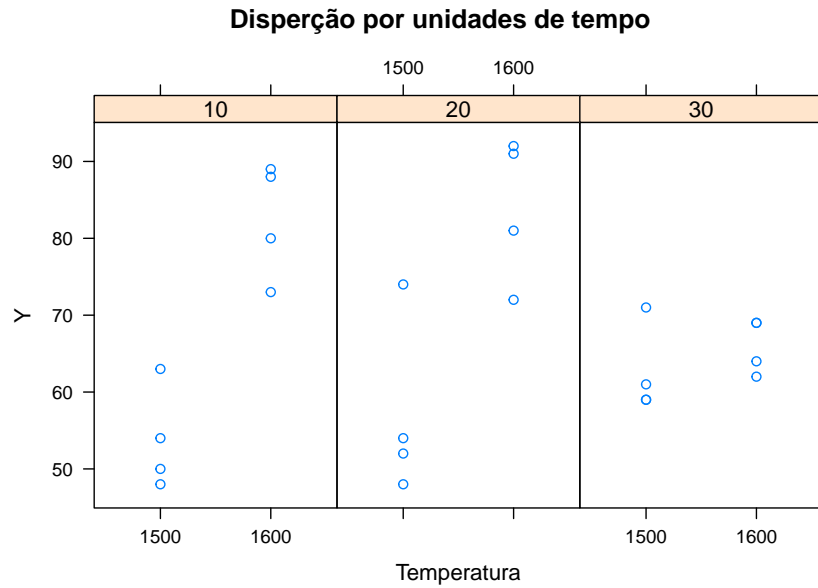
Pode-se olhar esse experimento como do tipo Split-Plot. E cada Plot possui 3 unidades experimentais: 10, 20 e 30 minutos.

- (b) Faça uma análise descritiva dos dados.

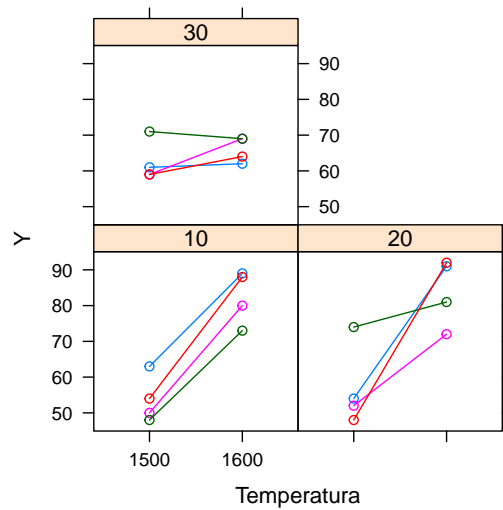
Resolução

Fazendo a análise descritiva, temos:





Interação entre temperatura e força de atração



Medidas Descritivas

Temperatura	Média	Desvio Padrão	Mediana
1500	57.75	8.49	56.6
1600	77.5	10.75	76.5

Tempo	Média	Desvio Padrão	Mediana
10	68.12	16.71	68
20	70.5	17.43	73
30	64.25	4.8	63

Rep	Média	Desvio Padrão	Mediana
1	70	15.82	62.5
2	63.67	11.91	64
3	69.33	11.22	72
4	67.5	18.26	61.5

Com a análise descritiva acima podemos notar que o valor mediano da temperatura de 1500 °F é menor do que o valor mediano da temperatura de 1600 °F, desconsiderando o tempo. Para cada tempo separadamente,

observamos que ao tempo de 10 minutos todas as replicas possuem valores menores para a temperatura de 1500 °F. Ao tempo de 20 minutos apenas uma réplica da temperatura de 1500 °F possui valor na força de tração do aço próximo aos valores de 1600 °F. Já, com 30 minutos de aquecimento os valores se apresentam próximos nas duas temperaturas.

- (c) Escreva a equação de um modelo apropriado para a análise dos dados. Não esqueça de descrever os fatores de interesse, seus respectivos níveis e a variável resposta.

Resolução

O modelo apropriado é:

$$y_{ijk} = \mu + \alpha_i + \rho_{k(i)} + \beta_j + (\alpha\beta)_{ij} + e_{ijk},$$

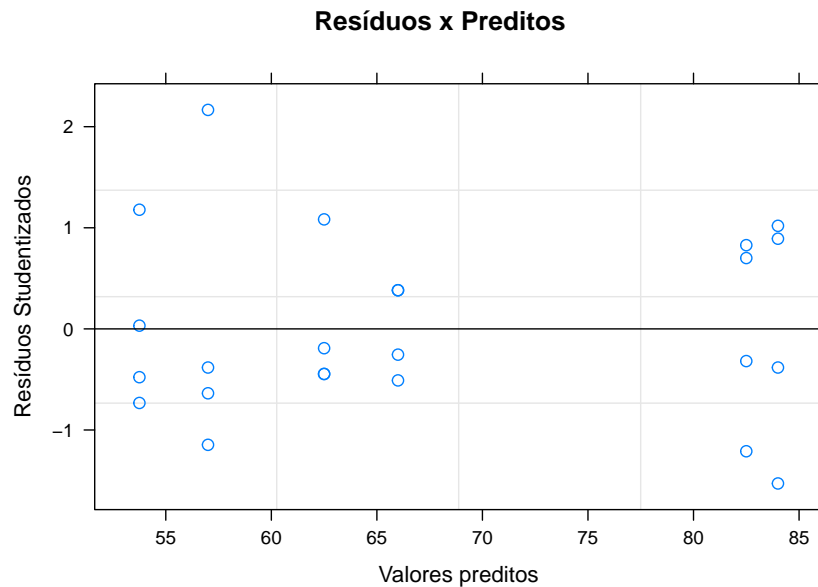
$$i = 1, 2, 3, 4; j = 1, 2; k = 1, 2, 3$$

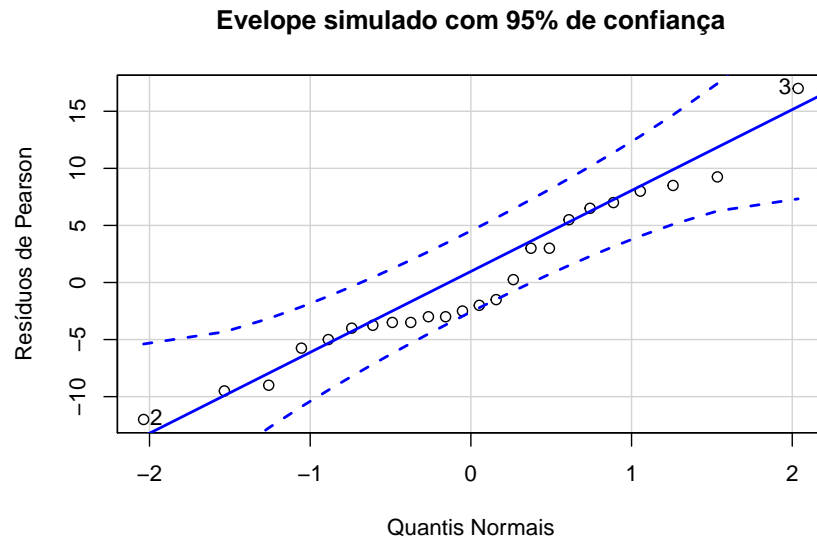
em que, μ é a média da variável resposta sob a casela de referência; α_i é o efeito fixo do nível i do fator de Repetição; β_j é o efeito fixo do nível j do fator Temperatura; $(\alpha\beta)_{ij}$ é o efeito de interação entre o nível i do fator de Repetição e o nível j do fator Temperatura e $\rho_{k(i)}$ é o efeito aleatório do tempo no nível k hierárquico ao nível i do fator de Repetição. y_{ijk} é a força de tração do aço na i -ésima Repetição, j -ésima Temperatura e k -ésimo tempo.

- (d) Ajuste o modelo proposto e faça uma análise de sua adequação.

Resolução

Observando o resíduo de Pearson, temos:





```
## 3 2
## 8 17
```

Verificamos então que os resíduos estão independentes e sem nenhum problema com a normalidade.

- (e) Teste as hipóteses de interesse, assumindo que os fatores temperatura e tempo de aquecimento têm efeitos fixos.

Resolução

```
## Linear mixed-effects model fit by REML
## Data: Dados2
##      AIC      BIC    logLik
## 149.5707 156.6936 -66.78533
##
## Random effects:
## Formula: ~1 | Rep
##      (Intercept) Residual
## StdDev: 0.0003982019  7.84839
##
## Fixed effects: Y ~ Temperatura * Tempo
##              Value Std.Error DF   t-value p-value
## (Intercept)    53.75   3.924195  15  13.697077  0.0000
## Temperatura1600    28.75   5.549650  15   5.180507  0.0001
## Tempo20             3.25   5.549650  15   0.585623  0.5668
## Tempo30             8.75   5.549650  15   1.576676  0.1357
## Temperatura1600:Tempo20  -1.75   7.848390  15  -0.222976  0.8266
## Temperatura1600:Tempo30 -25.25   7.848390  15  -3.217220  0.0058
## Correlation:
##              (Intr) Tm1600 Temp20 Temp30 T1600:T2
## Temperatura1600  -0.707
## Tempo20          -0.707  0.500
```

```
## Tempo30          -0.707  0.500  0.500
## Temperatura1600:Tempo20  0.500 -0.707 -0.707 -0.354
## Temperatura1600:Tempo30  0.500 -0.707 -0.354 -0.707  0.500
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -1.5289760 -0.4857684 -0.2866830  0.7326344  2.1660494
##
## Number of Observations: 24
## Number of Groups: 4
```

Table 2: Anova

	numDF	denDF	F-value	p-value
(Intercept)	1	15	1781.823	0.000
Temperatura	1	15	37.995	0.000
Tempo	2	15	1.293	0.303
Temperatura:Tempo	2	15	6.455	0.009

Em que podemos notar que existe efeito de interação e de temperatura a um nível de significância de 5%.

(f) Realize comparações múltiplas quando necessário.

Resolução

Como existe efeito no fator temperatura iremos realizar comparações múltiplas:

Table 3: Comparações múltiplas 1

	Estimate	lwr	upr
1600 - 1500	19.75	13.4701	26.0299

Em que podemos notar que o efeito da temperatura 1600 °F é maior na força de atração do aço do que a temperatura de 1500 °F.

Para o efeito de interação temos:

Table 4: Comparações múltiplas 2

	Estimate	lwr	upr
Temperatura1600:Tempo20 - Temperatura1600:Tempo30	-23.5	-38.88256	-8.117439

Em que a interação da temperatura 1600 °F com o Tempo 20 minutos é menor que a interação da temperatura 1600 °F com o Tempo 30 minutos.

Códigos

```
library(lattice)
library(car)
library(nlme)
library(lmerTest)
library(multcomp)
library(agricolae)

# Exercício 1
dados <- read.csv("dados.csv")
dados$Bloco <- as.factor(dados$Bloco)
dados$Formulacao <- as.factor(dados$Formulacao)
dados$Pia <- as.factor(dados$Pia)

attach(dados)

# analise descritiva
boxplot(Y~Formulacao,main = "Formulação x Reposta",xlab = "Formulação",data = dados)

with(dados, xyplot(Y ~ Formulacao | Pia,main="Dispersão fomulação por pia",xlab="Formulação"))

tapply(Y,Formulacao,mean)
tapply(Y,Formulacao,sd)
tapply(Y,Formulacao,median)

model <- lm(Y~Bloco+Formulacao+Pia)

round(Anova(model,type="III"),3)

BIBslide.aov <- aov(Y ~ Formulacao + Error(Bloco/Pia))
summary(BIBslide.aov)

BIB.test(Bloco, Formulacao, Y, test = c("tukey"),
          alpha = 0.05, group = TRUE,console=T)
detach(dados)

# Exercício 3

# item b
Dados2 <- read.csv("dados2.csv")
Dados2$Temperatura <- as.factor(Dados2$Temperatura)
Dados2$Tempo <- as.factor(Dados2$Tempo)
Dados2$Rep <- as.factor(Dados2$Rep)

attach(Dados2)
boxplot(Y~Temperatura, main="Temperatura x Força de Atração",
        ylab = "Força de Atração",data = Dados2)

with(Dados2, xyplot(Y ~ Temperatura | Tempo,
                    main="Dispersão por unidades de tempo"))

with(Dados2, xyplot(Y ~ Temperatura | Tempo, groups = Rep, aspect = "xy",
```

```

        type = "o",
        main="Interação entre temperatura e força de atração"))

tapply(Y,Temperatura,mean)
tapply(Y,Temperatura,sd)
tapply(Y,Temperatura,median)

tapply(Y,Tempo,mean)
tapply(Y,Tempo,sd)
tapply(Y,Tempo,median)

tapply(Y,Rep,mean)
tapply(Y,Rep,sd)
tapply(Y,Rep,median)

# item d

fit <- lme(Y ~ Temperatura*Tempo, random = ~1| Rep, data = Dados2)

plot(fit, main="Resíduos x Preditos",ylab="Resíduos Studentizados",
     xlab="Valores preditos")

qqPlot(residuals(fit),envelope=.95,xlab="Quantis Normais",
       ylab = "Resíduos de Pearson",
       main="Envelope simulado com 95% de confiança")

# item e
summary(fit)
round(anova(fit),3)

# item f

confint(glht(fit, linfct=mcp(Temperatura="Tukey",
                             interaction_average = T)))$confint

contrast.matrix <- rbind("Temperatura1600:Tempo20 - Temperatura1600:Tempo30" =
                          c(0, 0, 0, 0, -1, 1))

confint(glht(fit, contrast.matrix))$confint

detach(Dados2)

```