

# Lista 4 - MAE0317

*Guilherme Navarro N<sup>o</sup>USP:8943160 Leonardo Noronha N<sup>o</sup>USP:9793436*

## Exercício 1

Considere os dados do arquivo PULSE em que, na condição inicial, 92 estudantes foram avaliados segundo a pulsação (P1), hábito de fumar, gênero, altura, peso e atividade esportiva. Esses estudante jogaram uma moeda e aqueles que tiraram “cara” correram (Ran=1) um percurso e os demais permaneceram em repouso. A pulsação (P2) foi então novamente medida.

(a) Use o teste t para verificar se há efeito da corrida na pulsação P2. Faça suposições necessárias.

### Resolução

Supomos que há normalidade, independência e homocedasticidade nos resíduos, logo usaremos o teste para amostras independentes:

$$t_0 = \frac{\bar{Y}_1 - \bar{Y}_2}{Sc\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{(n_1+n_2-2)}$$

em que

$$Sc^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Aplicando aos dados da questão:

$$n_1 = 35 \quad n_2 = 57 \quad \bar{Y}_1 = 92.51429 \quad \bar{Y}_2 = 72.31579 \quad S_1 = 18.94321 \quad S_2 = 9.948363 \quad S_c^2 = 197.1451 \quad t_0 = 6.698901$$

Onde  $\bar{Y}_1$  é a média da pulsação P2 dos estudantes que correram e  $\bar{Y}_2$  é a média da pulsação P2 dos estudantes que não correram. Não rejeitamos a hipótese de a pulsação ser igual, a um nível de significância de 5% se:

$$t_{(0.025;n_1+n_2-2)} \leq t_0 \leq t_{(0.975;n_1+n_2-2)}$$

Como  $t(0.025;95) = -1.986675$  e  $t(0.975;95) = 1.986675$  e  $t_0 = 6.698901$ , rejeitamos a hipótese, ou seja, a média da pulsação de estudantes que correram não é igual a média da pulsação de estudantes que não correram

(b) Use ANOVA para verificar se há efeito da corrida na resposta P2. Compare os resultados com (a).

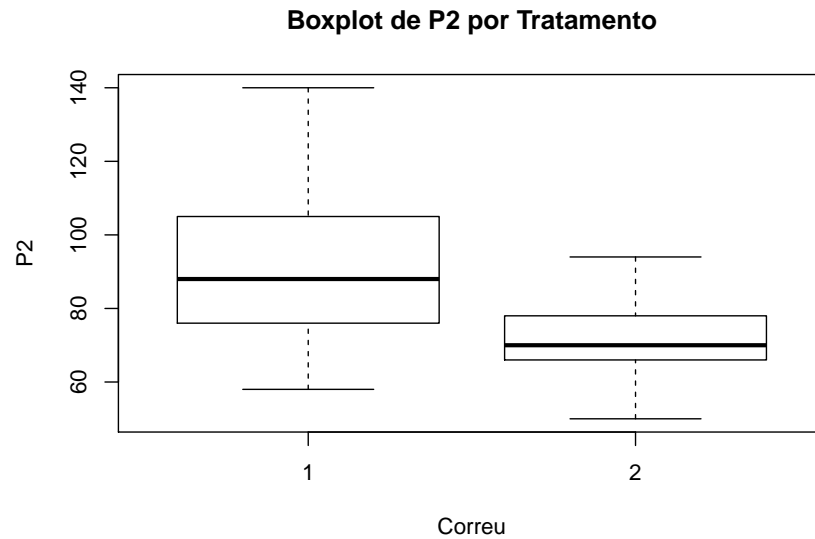
### Resolução

Foi criado um modelo do tipo:

$$Y_{ij} = \mu + \tau_j + e_{ij}$$

em que  $\mu$  é a média da pulsação P2 e o tratamento usado é se o estudante correu ou não. É feito então o boxplot, em que 1 são os alunos que correram e 2 os alunos que não correram:

Gráfico 1: Boxplot de P2 por Tratamento



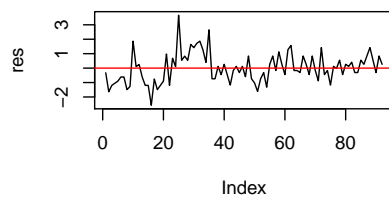
Aparentemente os valores observados em 2 possuem menor variabilidade mas não é possível afirmar que há diferença de tratamento. Realizando a tabela de ANOVA para o modelo, obtemos:

Table 1: Tabela de ANOVA

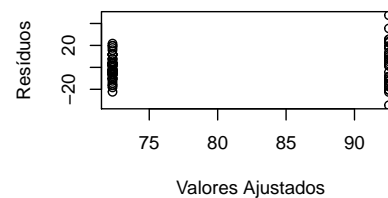
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Tratb	1	8846.941	8846.9414	44.87528	0
Residuals	90	17743.059	197.1451	NA	NA

Podemos verificar que não há diferença entre a pulsação P2 entre os estudantes que correram e não correram, resultado contrário ao do item anterior. É feita a análise de resíduos para o modelo:

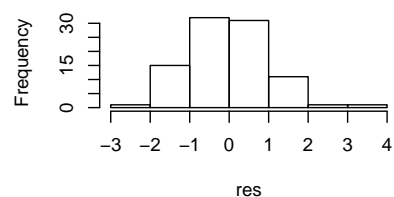
**Gráfico 2: Resíduos vs índices**



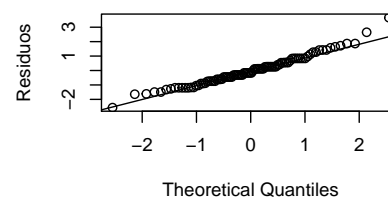
**Gráfico 3: Resíduos vs Preditos**



**Gráfico 4: Histograma dos resíduos**



**Gráfico 5: QQPlot Normal**



Aparentemente as suposições de normalidade e homocedasticidade estão comprometidas. Realizamos os teste de Shapiro-Wilk e de Levene:

```
##
## Shapiro-Wilk normality test
##
## data:  res
## W = 0.97523, p-value = 0.07681

## Loading required package: carData

## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value    Pr(>F)
## group 1  18.543 4.223e-05 ***
##      90
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## Bartlett test of homogeneity of variances
##
## data:  P2 by Tratb
## Bartlett's K-squared = 18.01, df = 1, p-value = 2.198e-05
```

Verificamos que realmente não há homocedasticidade mas existe uma normalidade fraca segundo o teste de Shapiro-Wilk.

- (c) Para cada grupo, que correu e para o em repouso, use o teste t para verificar se há efeito da corrida, isto é, para comparar as médias de P2 e P1. Faça as suposições necessárias.

## Resolução

Supomos que há normalidade, independência e homocedasticidade nos resíduos, logo usaremos o teste para amostras pareadas:

$$t_0 = \frac{\bar{d}}{\frac{S_d}{\sqrt{n}}} \sim t_{(n-1)}$$

Aplicando aos dados da questão, faremos dois testes:

$$n_1 = 35 \quad n_2 = 57 \quad \bar{d}_1 = 18.91429 \quad \bar{d}_2 = -0.1052632 \quad S_{d_1} = 15.04967 \quad S_{d_2} = 4.160547 \quad t_1 = 7.435276 \quad t_2 = -0.1910132$$

Não rejeitamos a hipótese de a pulsação ser igual, a um nível de significância de 5% se:

$$t_{(0.025; n_i-1)} \leq t_i \leq t_{(0.975; n_i-1)}$$

Para os estudantes que correram,  $t(0.025; 34) = -2.032245$  e  $t(0.975; 34) = 2.032245$  e  $t_1 = 7.435276$ , rejeitamos a hipótese, ou seja, a média da pulsação P2 dos estudantes que correram não é igual a média da pulsação P1.

Para os estudantes que não correram,  $t(0.025; 56) = -2.003241$  e  $t(0.975; 56) = 2.003241$  e  $t_2 = -0.1910132$ , não rejeitamos a hipótese, ou seja, a média da pulsação P2 dos estudantes não que correram é igual a média da pulsação P1.

(d) Calcule  $Dif = P2 - P1$  e use ANOVA para verificar se há efeito da corrida.

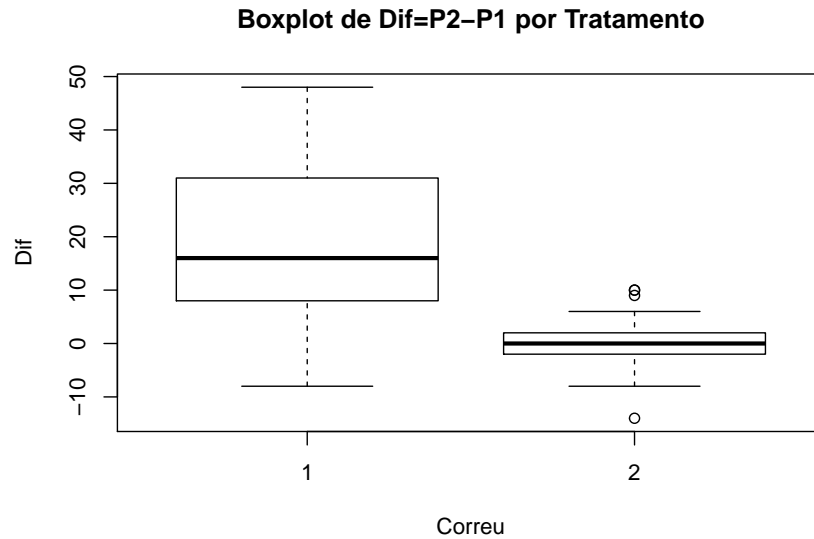
### Resolução

Foi criado um modelo do tipo:

$$Y_{ij} = \mu + \tau_j + e_{ij}$$

em que  $\mu$  é a média da diferença entre P2 e P1 e o tratamento usado é se o estudante correu ou não. É feito então o boxplot, em que 1 são os alunos que correram e 2 os alunos que não correram:

Gráfico 6: Boxplot de  $Dif = P2 - P1$  por Tratamento



Aparentemente os valores observados em 2 possuem menor variabilidade mas não é possível afirmar que há diferença de tratamento. Realizando a tabela de ANOVA para o modelo, obtemos:

Table 2: Tabela de ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Tratd	1	7844.324	7844.32350	81.42792	0
Residuals	90	8670.111	96.33457	NA	NA

Podemos verificar que não há diferença entre  $P2 - P1$  entre os estudantes que correram e não correram. É feito a análise de resíduos para o modelo:

Gráfico 7: Resíduos vs índices

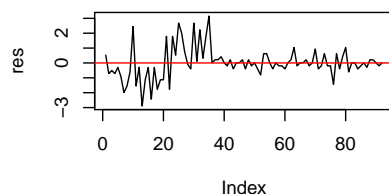


Gráfico 8: Resíduos vs Preditos

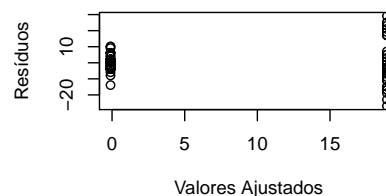


Gráfico 9: Histograma dos resíduos

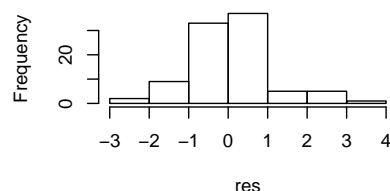
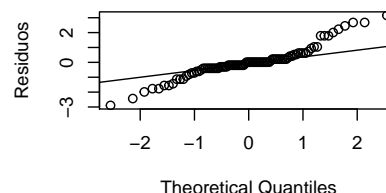


Gráfico 10: QQPPlot Normal



Aparentemente as suposições de normalidade e homocedasticidade estão comprometidas. Realizamos os teste de Shapiro-Wilk e de Levene:

```
##
## Shapiro-Wilk normality test
##
## data:  res
## W = 0.91854, p-value = 2.501e-05

## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value    Pr(>F)
## group 1  43.449 2.882e-09 ***
##      90
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Verificamos que realmente não há normalidade e homocedasticidade.

- (e) Considere um delineamento fatorial 2x2 com dois fatores cruzados, Tempo (inicial e final) e Corrida (sim e não). Ainda, considere indivíduo como fator bloco. Neste caso, use ANOVA para verificar se há efeito da corrida na pulsação. Que suposições foram feitas? Compare os resultados com (c) e (d).

## Resolução

Foi criado um modelo do tipo:

$$Y_{ijkl} = \mu + \tau_j + \beta_i + \gamma_l + e_{ijkl}$$

em que  $\mu$  é a média de todas as pulsações, os tratamentos usados são se o estudante correu ou não, se a pulsação é do tempo inicial ou final, a interação entre esses tratamentos e o bloco são os indivíduos. São feito boxplots, em que no primeiro é a relação entre os estudantes que correram(1) e os estudantes que não correram(2), já no segundo, é a relação entre os tempos inicial e final.

Gráfico 11: Boxplot das Medidas por Tratamento se Correu ou não

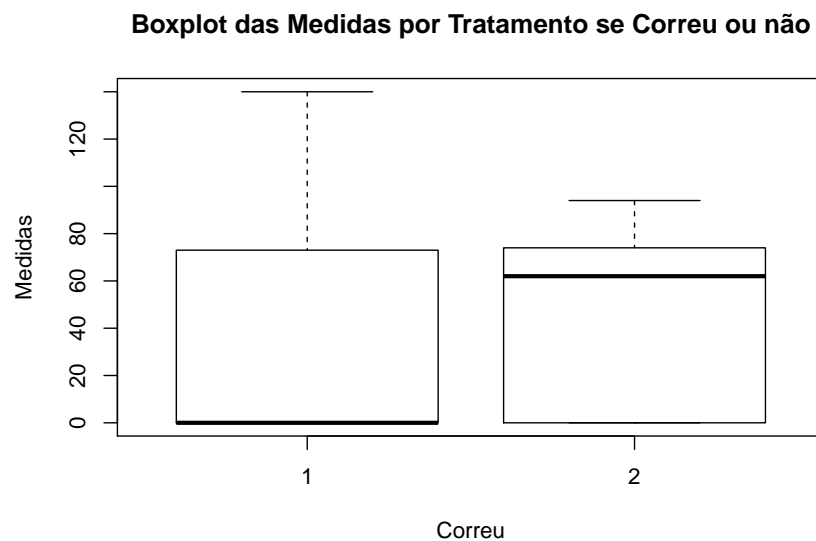
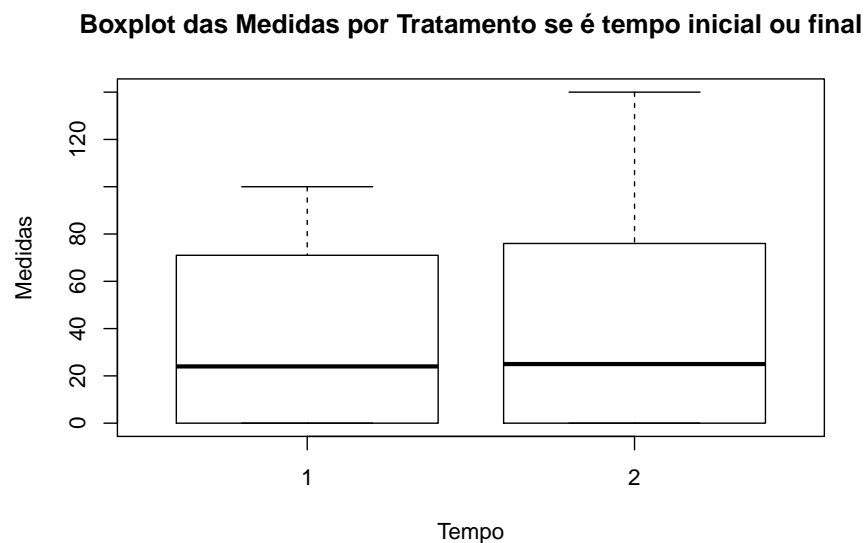


Gráfico 12: Boxplot das Medidas por Tratamento se é tempo inicial ou final

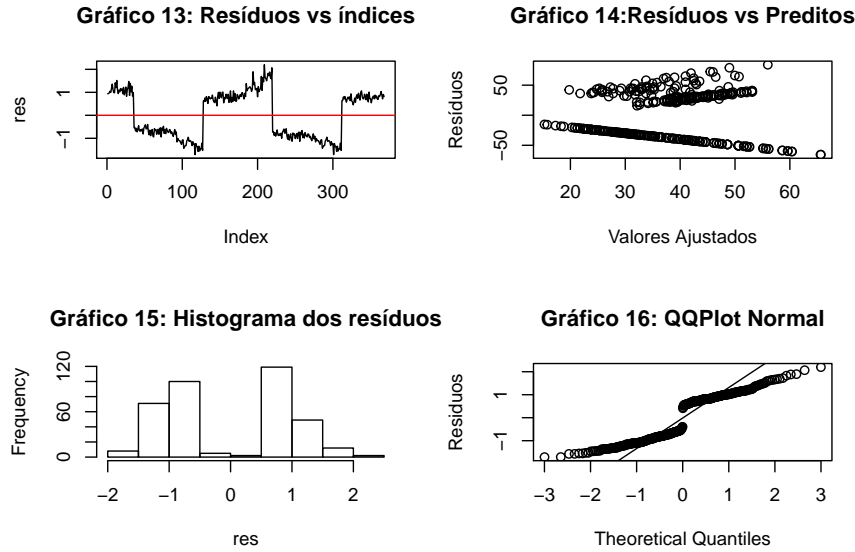


Aparentemente os valores dos estudantes que não correram tem uma variabilidade menor, assim como no tempo inicial e não é possível dizer nada sobre as médias. Realizando a tabela de ANOVA para o modelo, obtemos:

Table 3: Tabela de ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
correu_ou_ncorreu	1	16125.261	16125.2609	8.0884248	0.0047918
inicial_ou_final	1	1169.391	1169.3913	0.5865662	0.4444117
Bloco	91	14680.609	161.3254	0.0809207	1.0000000
correu_ou_ncorreu:inicial_ou_final	1	1212.565	1212.5652	0.6082223	0.4361336
Residuals	273	544258.783	1993.6219	NA	NA

Podemos verificar que não há diferença entre os estudantes que correram e não correram mas há uma diferença entre o tempo inicial e final, assim como a interação. É feito a análise de resíduos para o modelo:



Aparentemente as suposições de normalidade e homocedasticidade estão muito comprometidas, chegando a ser desnecessário os testes de Shapiro-Walk, Levene e Bartlett.

- (f) Considere um delineamento em que a pulsação (inicial e final) é hierárquica dentro de corrida (sim e não). Adote o modelo de efeitos fixos dos fatores e indivíduo como um fator Bloco. Há efeito de corrida? Compare os resultados com (e).

## Resolução

Foi criado um modelo do tipo:

$$Y_{ijkl} = \mu + \tau_j + \beta_{i(j)} + \gamma_l + e_{ijkl}$$

em que  $\mu$  é a média entre todas as pulsações, os tratamentos usados são se o estudante correu ou não, hierárquicamente se a pulsação é do tempo inicial ou final e o bloco são os indivíduos.

Gráfico 17: Boxplot das Medidas por Tratamento se Correu ou não

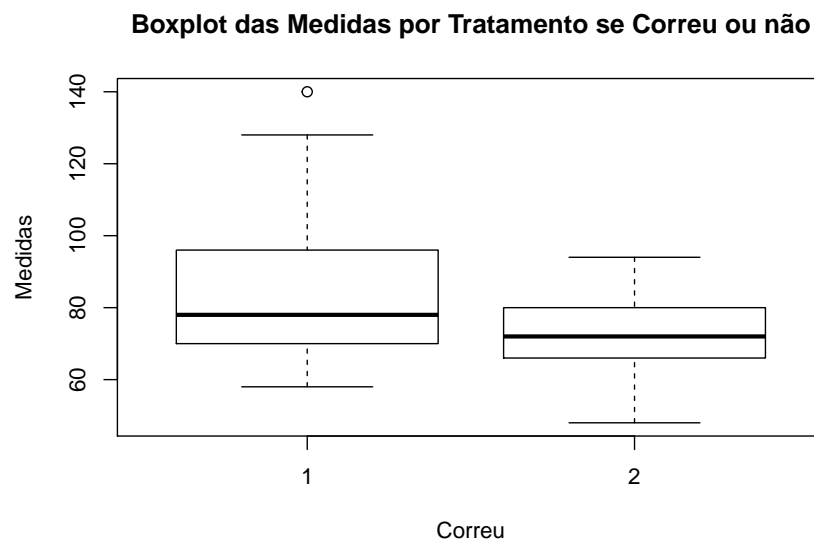
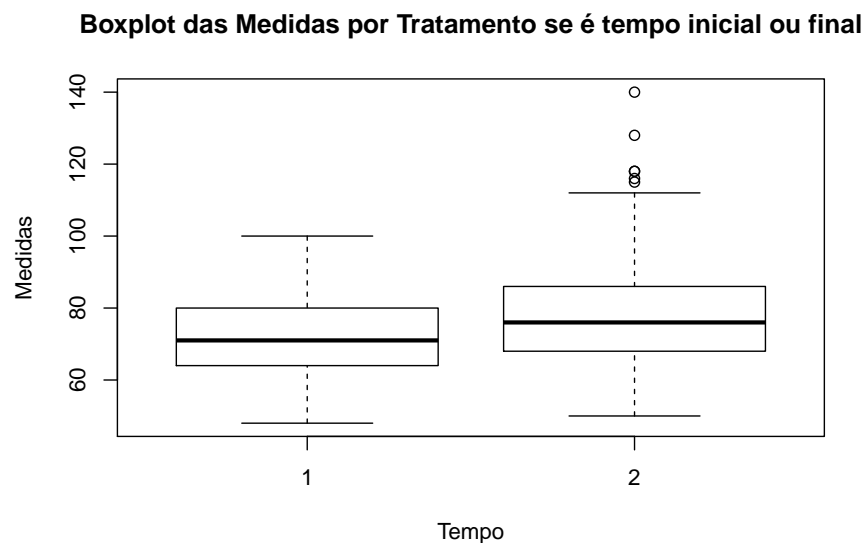


Gráfico 18: Boxplot das Medidas por Tratamento se é tempo inicial ou final



Aparentemente os valores dos estudantes que não correram tem uma variabilidade menor, assim como no tempo inicial e não é possível dizer nada sobre as médias. Realizando a tabela de ANOVA para o modelo, obtemos:

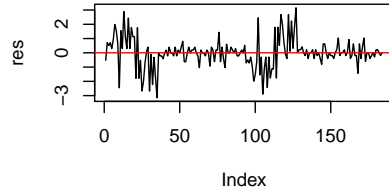
Table 4: Tabela de ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
correu_ou_ncorreu	1	4954.920	4954.91965	102.868984	0
Bloco	90	24406.298	271.18109	5.629985	0
correu_ou_ncorreu:inicial_ou_final	2	6260.944	3130.47218	64.991668	0
Residuals	90	4335.056	48.16728	NA	NA

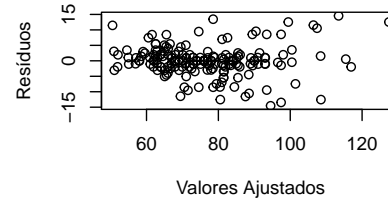


Podemos verificar que não há diferença entre os estudantes que correram e não correram e entre o tempo inicial e final hierárquicamente. É feito a análise de resíduos para o modelo:

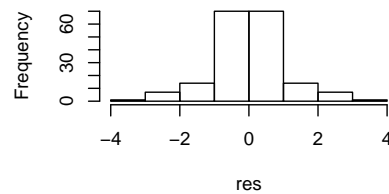
**Gráfico 19: Resíduos vs índices**



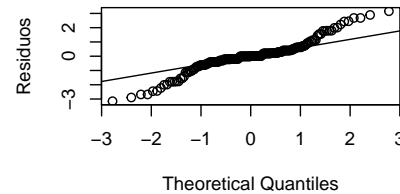
**Gráfico 20: Resíduos vs Preditos**



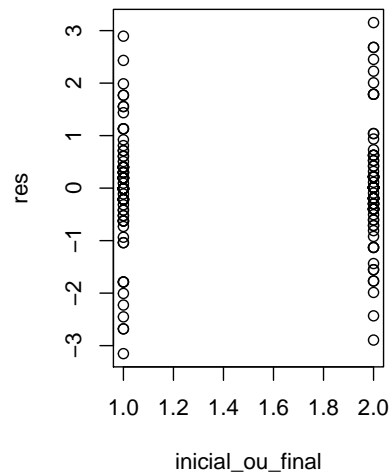
**Gráfico 21: Histograma dos resíduos**



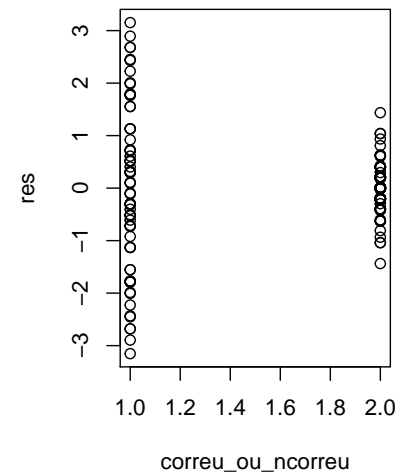
**Gráfico 22: QQPlot Normal**



**Gráfico 23: Resíduos vs B(A)**



**Gráfico 24: Resíduos vs A**



Aparentemente as suposições de normalidade e homocedasticidade estão comprometidas. Realizamos os teste de Shapiro-Wilk e de Levene:

```
##
##  Shapiro-Wilk normality test
##
## data:  res
## W = 0.93009, p-value = 9.644e-08

## Levene's Test for Homogeneity of Variance (center = median)
```

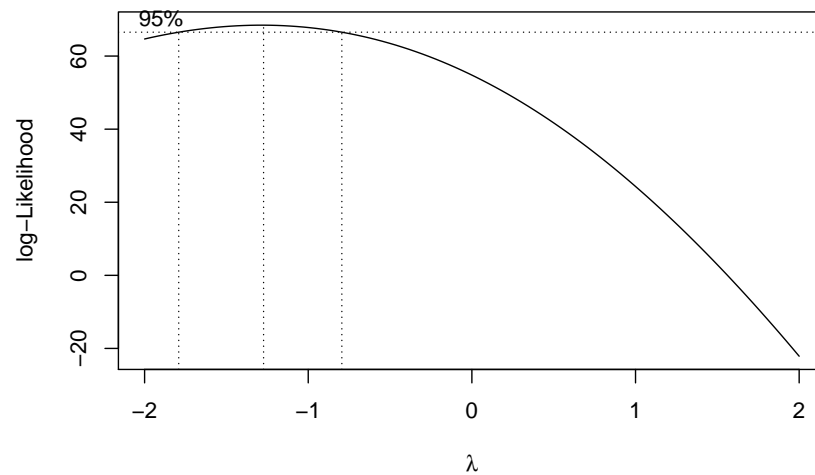
```
##           Df F value   Pr(>F)
## group     3  8.5545 2.43e-05 ***
##           180
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Verificamos que realmente não há normalidade e homocedasticidade.

(g) Finalmente, qual modelo você recomendaria na análise desses dados? Justifique.

## Resolução

Ao analisar os modelos podemos ver que o modelo do item f é o mais completo, entretanto é necessário uma transformação de boxcox:



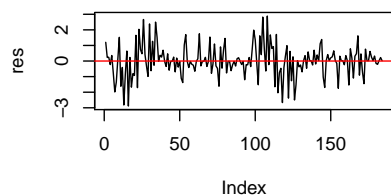
Faremos a transformação  $Medidas^{-1}$ . Realizando novamente a tabela de ANOVA:

Table 5: Tabela de ANOVA

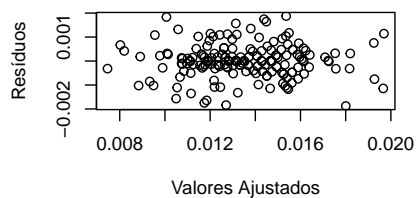
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
correu_ou_ncorreu	1	0.0001034	0.0001034	109.861792	0
Bloco	90	0.0007087	0.0000079	8.363382	0
correu_ou_ncorreu:inicial_ou_final	2	0.0001220	0.0000610	64.811596	0
Residuals	90	0.0000847	0.0000009	NA	NA

Concluimos que não há em diferença entre as médias. Faremos a análise de resíduos:

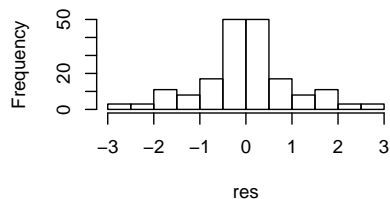
**Gráfico 24: Resíduos vs índices**



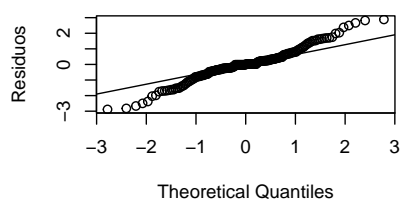
**Gráfico 25: Resíduos vs Preditos**



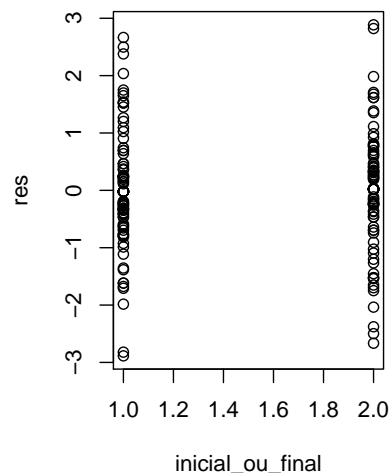
**Gráfico 26: Histograma dos resíduos**



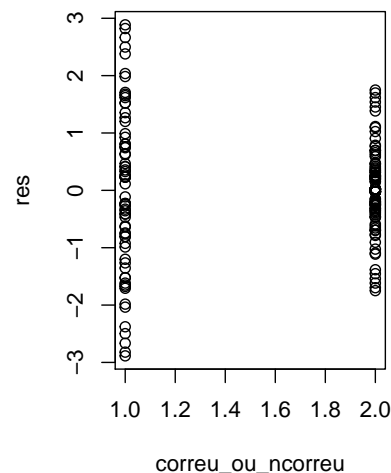
**Gráfico 27: QQPlot Normal**



**Gráfico 28: Resíduos vs B(A)**



**Gráfico 29: Resíduos vs A**



Aparentemente a variância foi arrumada mas a normalidade continua comprometida, faremos agora então os testes de Shapiro-Walk e Levene:

```
##
## Shapiro-Wilk normality test
##
## data:  res
## W = 0.97046, p-value = 0.0006146

## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 3  0.5538 0.6462
##      180
```

Realmente a normalidade, apesar de melhorar bastante, continua comprometida mas pelo fato de haver aleatorização não é preciso se preocupar tanto com essa premissa. Entretanto com a transformação de variável a homocedasticidade está presente no novo modelo.

### Exercício 3

Um estudo foi conduzido para avaliar o efeito de interação entre dois medicamentos (X e Y) usados para estimular o crescimento de crianças com uma particular síndrome que atinge o desenvolvimento infantil. Sabe-se que o efeito de cada medicamento é modesto, mas o efeito da combinação das duas drogas (X e Y) não é ainda conhecido. Os seguintes resultados foram obtidos da avaliação da taxa de crescimento de 16 pacientes pertencentes a quatro faixas etárias consideradas no estudo:

Faixa Etária	Pacientes			
	1	2	3	4
1	0.02 (A)	0.15 (B)	0.45 (D)	0.18 (C)
2	0.27 (B)	0.24 (C)	-0.01 (A)	0.58 (D)
3	0.11 (C)	0.35 (D)	0.14 (B)	-0.03 (A)
4	0.48 (D)	0.04 (A)	0.18 (C)	0.22 (B)

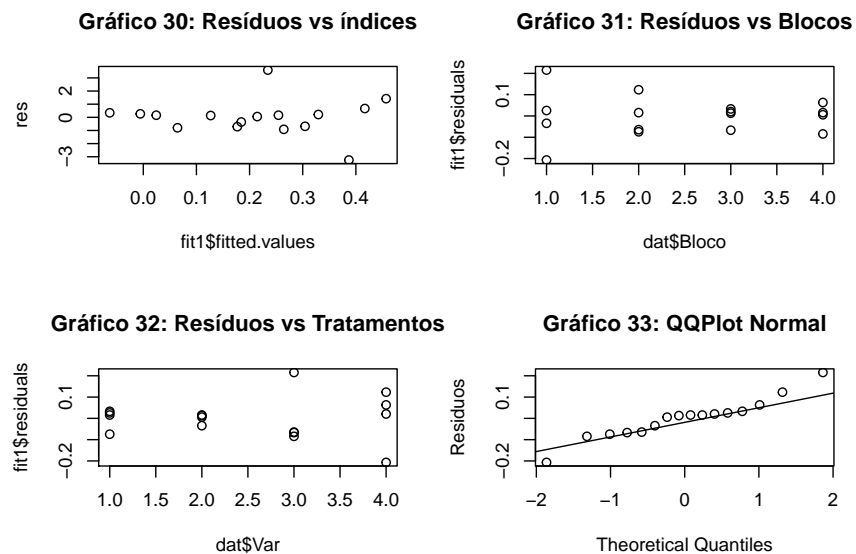
A=Placebo B= Droga X C=Droga Y D=Drogas X e Y

- (a) Ajuste um modelo de ANOVA para estes dados considerando um delineamento aleatorizado em blocos completos (DABC). Apresente e discuta os resultados obtidos

### Resolução

Table 6: Tabela de ANOVA

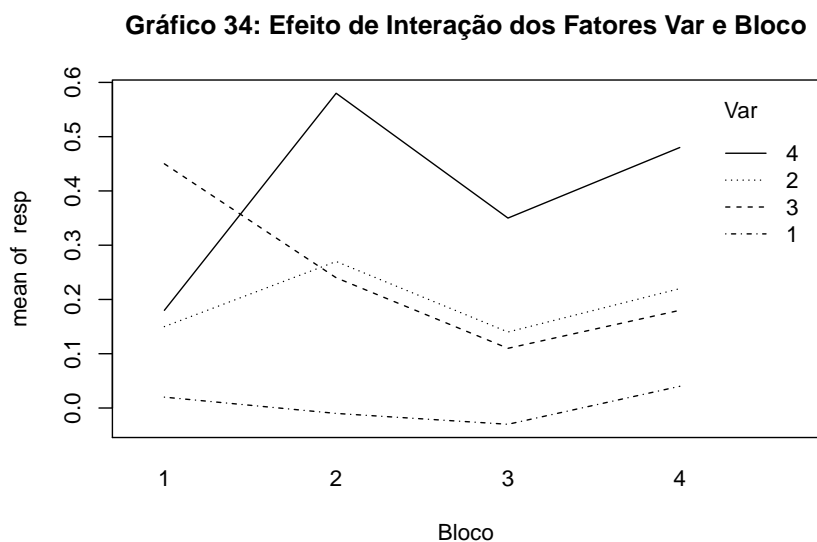
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Var	3	0.3145187	0.1048396	7.054297	0.0097337
Bloco	3	0.0346187	0.0115396	0.776459	0.5360128
Residuals	9	0.1337562	0.0148618	NA	NA



De acordo com a tabela de ANOVA com delineamento aleatorizado com blocos completos, onde a suposição de normalidade, homocedasticidade e independência entre observações podem ser confirmadas pela análise de resíduos acima, além disso temos que as drogas tem efeito de tratamento a um nível de significância fixado de 5%.

(b) Desenhe o gráfico de perfis de médias para estudar o efeito de interação entre as drogas X e Y.

### Resolução



Com o gráfico de interação podemos notar que há interação entre os blocos e os tratamentos, assim como o esperado.

- (c) Apresente contrastes apropriados entre as médias dos quatro tratamentos (A, B, C e D) que possam ser usados para testar os efeitos principais e de efeito de interação entre as drogas X e Y. Para estes contrastes obtenha os correspondentes intervalos de confiança. Como estes resultados podem ser usados para representar os 3 graus de liberdade do efeito de tratamento obtido em a)?

### Resolução

- (d) Como os medicamentos X e Y devem ser administrados para se obter eficiência máxima no tratamento de estímulo do crescimento em pacientes com tal síndrome?

### Resolução

- (e) Realize o teste de Tukey com 1 grau de liberdade para verificar se há efeito de interação entre medicamento e faixa etária.

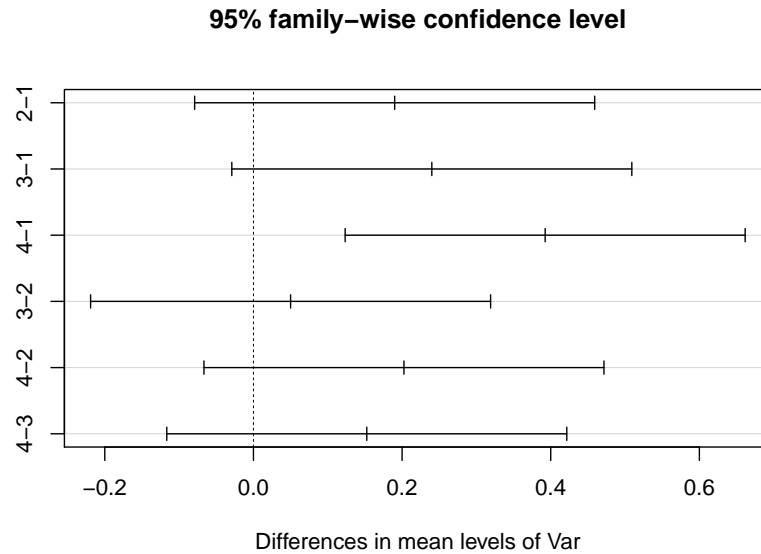
### Resolução

Table 7: Tabela de ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Var	3	0.3145187	0.1048396	6.2787135	0.0169514
Bloco	3	0.0346187	0.0115396	0.6910914	0.5826513
trbl	1	0.0001753	0.0001753	0.0104970	0.9209172
Residuals	8	0.1335810	0.0166976	NA	NA

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = resp ~ Var + Bloco)
##
## $Var
##      diff      lwr      upr      p adj
## 2-1 0.1900 -0.07910743 0.4591074 0.1937531
## 3-1 0.2400 -0.02910743 0.5091074 0.0830435
## 4-1 0.3925 0.12339257 0.6616074 0.0061828
## 3-2 0.0500 -0.21910743 0.3191074 0.9356157
## 4-2 0.2025 -0.06660743 0.4716074 0.1576020
## 4-3 0.1525 -0.11660743 0.4216074 0.3464902
```

Gráfico 35: Tukey efeito de Tratamento



Fazendo o teste de tukey com 1 grau de liberdade, a um nível de significância fixado de 5%, podemos aceitar o modelo aditivo, ou seja, existe de fato uma interação linear entre os tratamentos e blocos.

## Exercício 4

Considere a seguir os dados de dois experimentos conduzidos sob a estrutura de aleatorização de blocos completos. Em cada caso obtenha a tabela de ANOVA. Construa o gráfico de perfis de resposta de acordo com tratamento e bloco. Com base neste gráfico há indicação de efeito de interação entre os fatores tratamento e bloco? Realize o teste de Tukey com 1 grau de liberdade para verificar a existência desse efeito de interação. Faça recomendações aos pesquisadores dos dois experimentos.

Exp. 1	Bloco				
	1	2	3	4	
Tratamento	1	2	6	8	10
	2	3	6	12	15
	3	4	8	13	20

Exp. 2	Bloco				
	1	2	3	4	
Tratamento	1	2	6	8	10
	2	15	12	6	3
	3	16	13	7	4

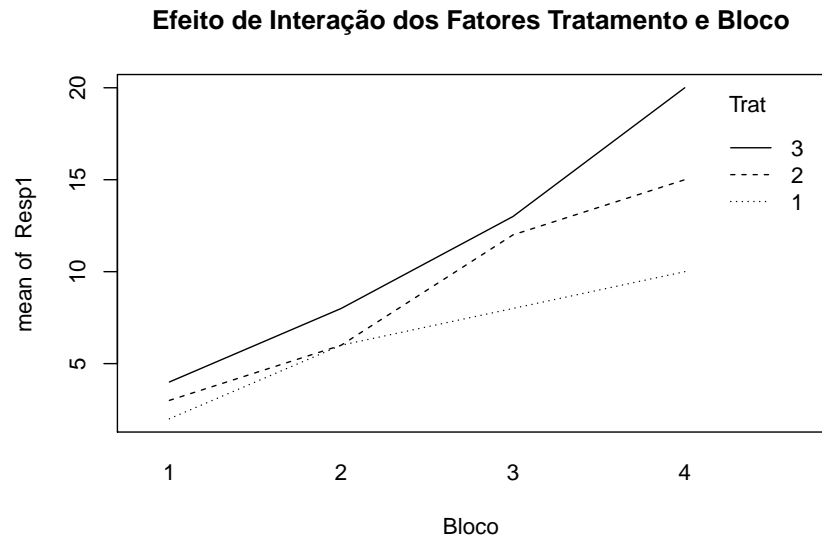
## Resolução

Considerando o Experimento 1, temos:

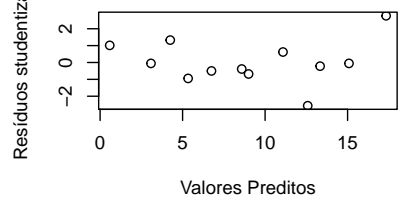
Table 8: Tabela de ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Trat	2	45.16667	22.583333	5.765957	0.0400835
Bloco	3	244.25000	81.416667	20.787234	0.0014294
Residuals	6	23.50000	3.916667	NA	NA

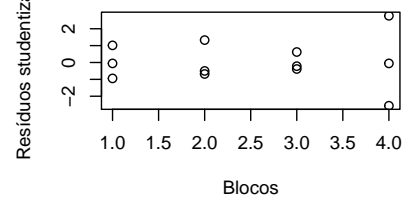
Gráfico 36: Interação entre os fatores tratamento e bloco



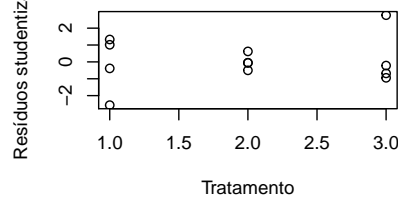
**Gráfico 37: Resíduos vs Preditos**



**Gráfico 38: Resíduos vs Blocos**



**Gráfico 39: Resíduos vs Tratamento**



**Gráfico 40: Resíduos vs índice**

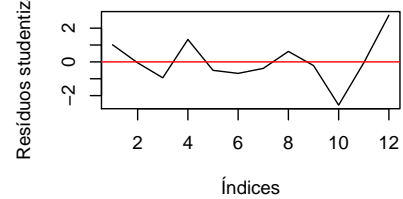




Gráfico 41: QPlot Normal

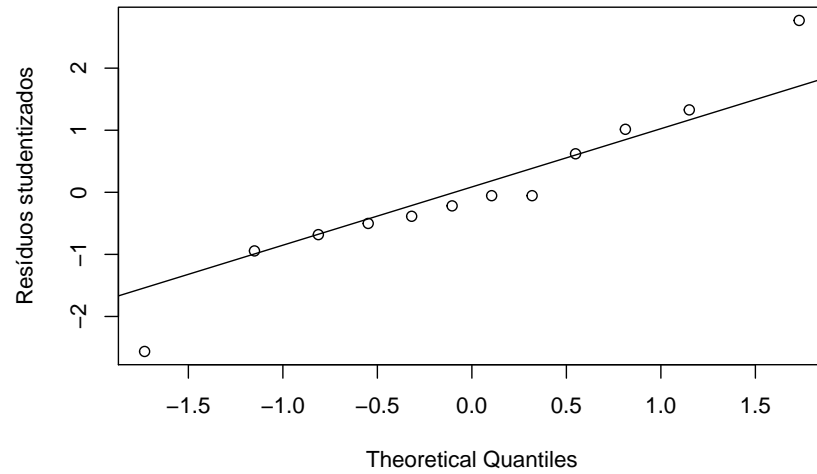
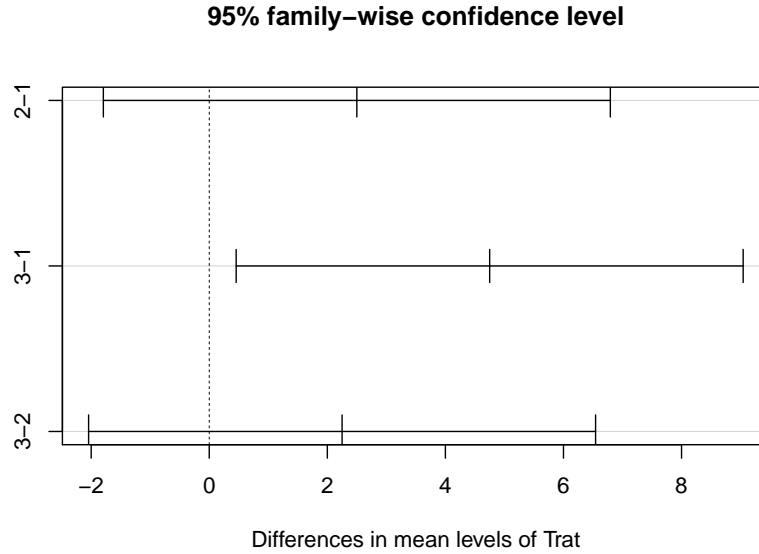


Table 9: Tabela de ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Trat	2	45.166667	22.5833333	23.19213	0.0029536
Bloco	3	244.250000	81.4166667	83.61149	0.0001080
trbl	1	18.631251	18.6312513	19.13351	0.0071937
Residuals	5	4.868749	0.9737497	NA	NA
Gráfico 42: Tukey	efeito de tr	atamento			

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = Resp1 ~ Trat + Bloco)
##
## $Trat
##      diff      lwr      upr    p adj
## 2-1 2.50 -1.7937575 6.793757 0.2519349
## 3-1 4.75  0.4562425 9.043757 0.0336705
## 3-2 2.25 -2.0437575 6.543757 0.3130070
```



Para o experimento 1 adotando o modelo de ANOVA com delineamento aleatorizado com blocos complexos (DABC) como podemos observar na tabela 8, temos algum problema com os blocos pois o mesmo não deveria ter efeito significativo, e também pelo gráfico 36 de interações podemos notar um cruzamento das retas do tratamento 1 com 2, nos sugerindo que existe interações entre o fator de tratamento e de bloco, fazendo o teste de Tukey com 1 grau de liberdade, a um nível de significância fixado de 5%, podemos aceitar o modelo aditivo, ou seja, existe de fato uma interação linear entre os tratamentos e blocos. Como recomendação ao pesquisador deve-se verificar se não existe nenhum fator que possa estar influenciando na coleta dos dados, ou até mesmo adotar um modelo de ANOVA com delineamento completamente aleatorizado com 1 fator em  $k$ -níveis.

Considerando o Experimento 2, temos:

Table 10: Tabela de ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Trat	2	26.00000	13.00000	0.5021459	0.6285807
Bloco	3	59.66667	19.88889	0.7682403	0.5522803
Residuals	6	155.33333	25.88889	NA	NA

Gráfico 43: Intereção entre os fatores tratamento e bloco

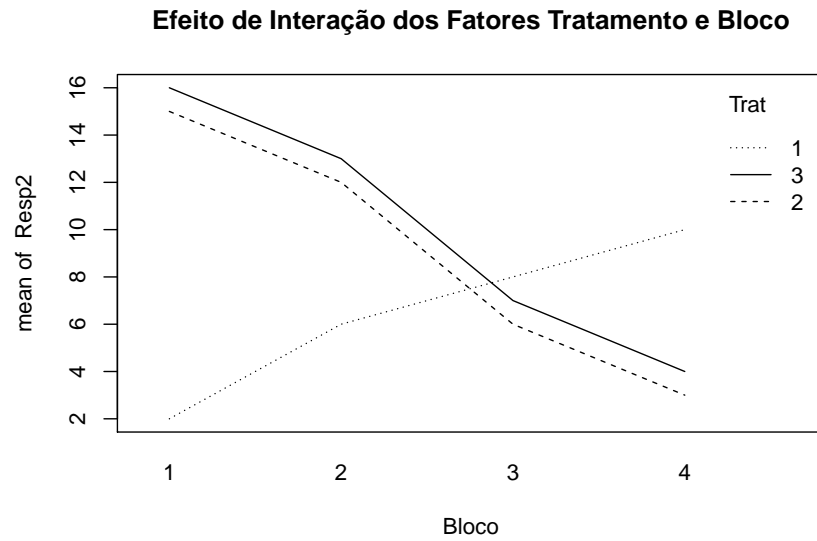


Gráfico 44: Resíduos vs Preditos

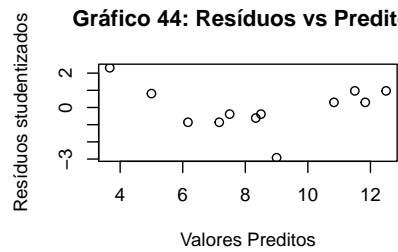


Gráfico 45: Resíduos vs Blocos

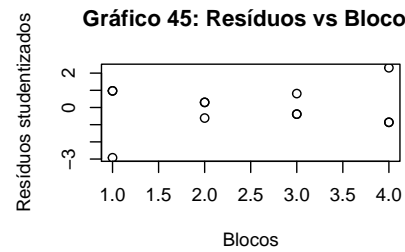


Gráfico 46: Resíduos vs Tratamento

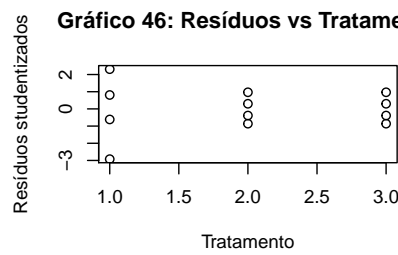


Gráfico 47: Resíduos vs índice

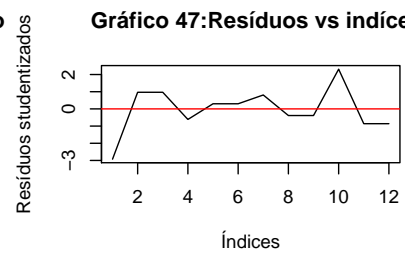


Gráfico 48: QQPlot Normal

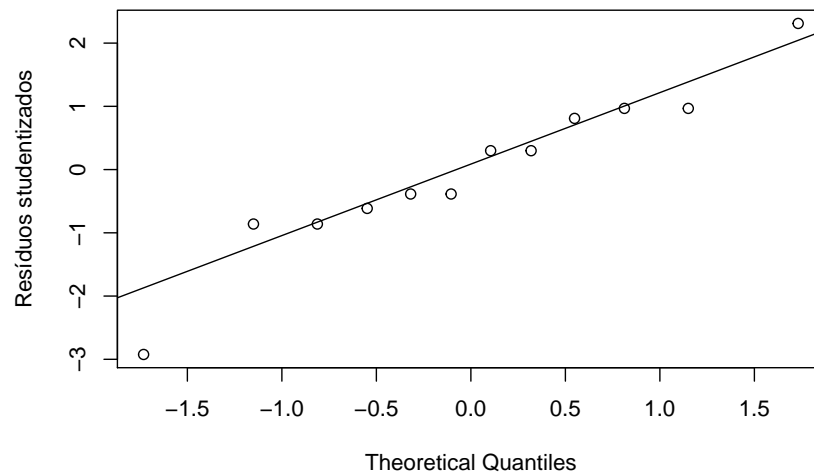


Table 11: Tabela de ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Trat	2	26.00000	13.000000	3.387368	0.1175025
Bloco	3	59.66667	19.888889	5.182383	0.0540636
trbl2	1	136.14439	136.144392	35.474701	0.0019078
Residuals	5	19.18894	3.837788	NA	NA

Para o experimento 1 adotando o modelo de ANOVA com delineamento aleatorizado com blocos complexos (DABC) como podemos observar na tabela 10, pelo gráfico 43 de interações podemos notar um cruzamento das retas do tratamento 1 com 2 e 3, nos sugerindo que existe interações entre o fator de tratamento e de bloco, fazendo o teste de tukey com 1 grau de liberdade, a um nível de significância fixado de 5%, rejeitamos o modelo aditivo, ou seja, não existe de fato uma interação linear entre os tratamentos e blocos, e como na análise descritiva nos sugeria uma interação, acreditamos que a mesma possa existir mas de uma forma não linear. Como recomendação ao pesquisador deve-se verificar se não houve nenhum erro na coleta dos dados, pois a relação bloco/tratamento 1, parece estar invertida.

## Exercício 5

Um engenheiro de produção estudou os efeitos do tipo de máquina e de operador na qualidade do processo de engarrafamento de bebidas em uma indústria. Três tipos diferentes de máquinas são usados no processo industrial de engarrafamento. Para operar cada máquina é preciso passar por um treinamento específico. Durante 5 dias de acompanhamento do processo de produção, foi contado o número de engarrafamentos dentro do padrão de qualidade exigido pelo órgão certificador. Os dados da porcentagem de garrafas aprovadas estão apresentados a seguir:

Máquina	1				2				3			
Operador	1	2	3	4	5	6	7	8	9	10	11	12
Dia 1	65	68	56	45	74	69	52	73	69	63	81	67
2	58	62	65	56	81	76	56	78	83	70	72	79
3	63	75	58	54	76	80	62	83	74	72	73	73
4	57	64	70	48	80	78	58	75	78	68	76	77
5	66	70	64	60	68	73	51	76	80	75	70	71

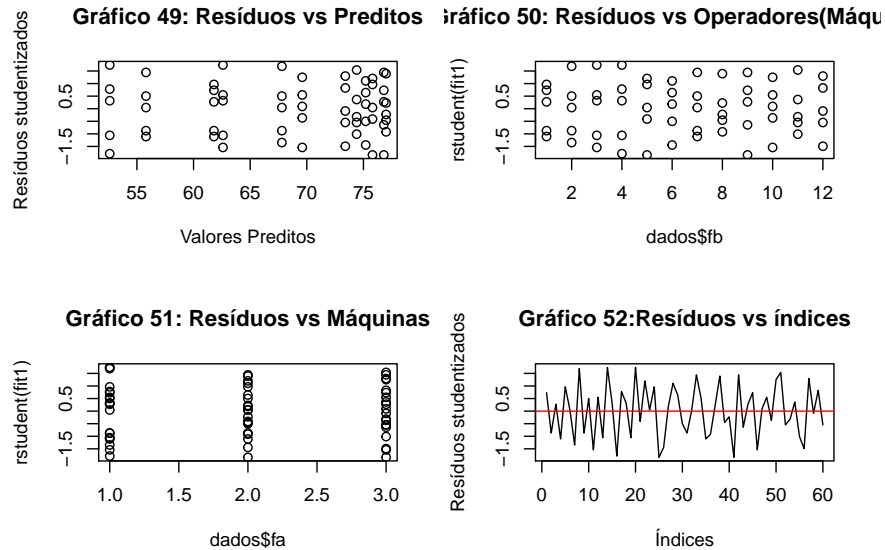
- (a) Ajuste um modelo de análise de variância com estrutura hierárquica dos fatores (assumidos fixos). Interprete os resultados. Faça uma análise de resíduos.

### Resolução

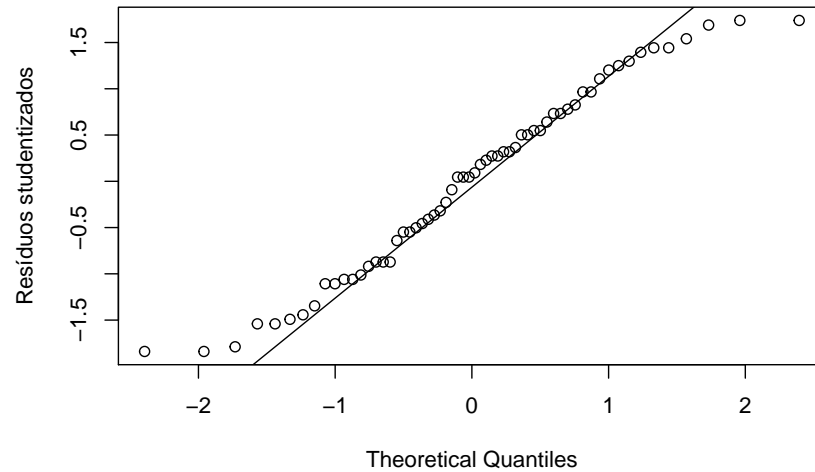
```
## The following object is masked _by_ .GlobalEnv:
##
##      resp
```

Table 12: Tabela de ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
fa	2	1695.633	847.8167	35.92444	0
fa:fb	9	2272.300	252.4778	10.69821	0
Residuals	48	1132.800	23.6000	NA	NA



**Gráfico 53: QQPlot Normal**



Com a tabela de ANOVA X com delineamento fatorial hierárquico, podemos concluir que há interesse em testar o efeito principal do operador em de cada máquina e o efeito principal das máquinas no engarrafamento de de bebidas em uma indústria, e com os resultados conseguimos notar que há efeito de interação, pois o *p-value* dos efeitos principais não praticamente nulos. Agora fazendo uma análise de resíduos para verificar se as conclusões da tabela de ANOVA são significantes estatisticamente, quanto a suposição de normalidade pelo gráfico X está tudo certo, quanto a homocedasticidade e independência pelos gráficos X e X também podemos concluir que está certo.

(b) Os efeitos de operador podem ser distinguidos dos efeitos de interação neste estudo? Justifique.

### Resolução

Sim, pois como estamos sob um estudo de um delineamento fatorial hierárquico, os efeitos de operador só podem ser comparados na mesma máquina.

(c) Realize comparações múltiplas entre as respostas médias dos operadores dentro de cada máquina.

### Resolução

Análise para máquina 1

```
## X.1.2. dif      li      ls sig
## 1  1-2 -6 -12.17759 0.1775914 ns
```

```
## X.1.3. dif      li      ls sig
## 1  1-3 -0.8 -6.977591 5.377591 ns
```

```
## X.1.3. dif      li      ls sig
## 1  1-3 9.2 3.022409 15.37759 *
```

```
## X.2.3. dif      li      ls sig
## 1  2-3 5.2 -0.9775914 11.37759 ns
```

```
## X.2.4. dif      li      ls sig
## 1    2-4 15.2 9.022409 21.37759 *
```

```
## X.3.4. dif      li      ls sig
## 1    3-4 10 3.822409 16.17759 *
```

Análise para máquina 2

```
## X.5.6. dif      li      ls sig
## 1    5-6 0.6 -5.577591 6.777591 ns
```

```
## X.5.7. dif      li      ls sig
## 1    5-7 20 13.82241 26.17759 *
```

```
## X.5.8. dif      li      ls sig
## 1    5-8 -1.2 -7.377591 4.977591 ns
```

```
## X.6.7. dif      li      ls sig
## 1    6-7 19.4 13.22241 25.57759 *
```

```
## X.6.8. dif      li      ls sig
## 1    6-8 -1.8 -7.977591 4.377591 ns
```

```
## X.7.8. dif      li      ls sig
## 1    7-8 -21.2 -27.37759 -15.02241 *
```

Análise para máquina 3

```
## X.9.10. dif      li      ls sig
## 1    9-10 7.2 1.022409 13.37759 *
```

```
## X.9.11. dif      li      ls sig
## 1    9-11 2.4 -3.777591 8.577591 ns
```

```
## X.9.12. dif      li      ls sig
## 1    9-12 3.4 -2.777591 9.577591 ns
```

```
## X.10.11. dif      li      ls sig
## 1    10-11 -4.8 -10.97759 1.377591 ns
```

```
## X.10.12. dif      li      ls sig
## 1    10-12 -3.8 -9.977591 2.377591 ns
```

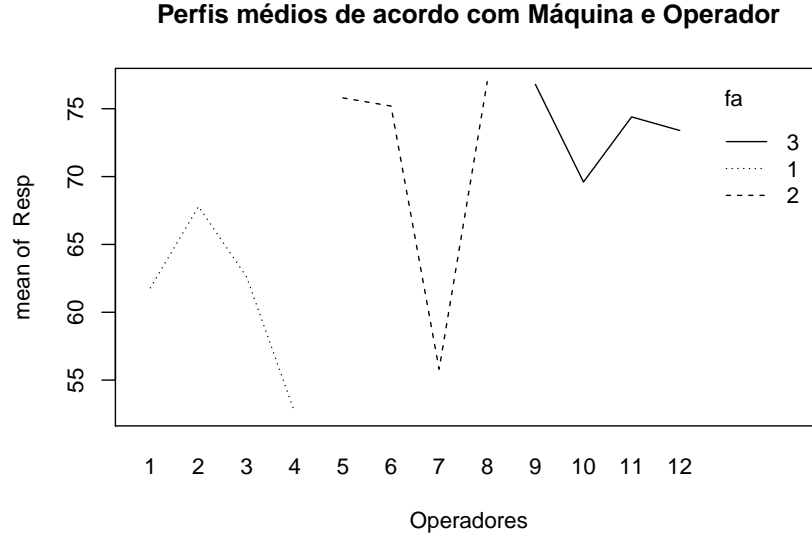
```
## X.11.12. dif      li      ls sig
## 1    11-12 1 -5.177591 7.177591 ns
```

Legenda: a um nível de significância de 5%, “\*” é significante e “ns” é não significante.

(d) Construa um gráfico com os perfis médios de resposta de acordo com Máquina e Operador. Nas condições do experimento, como o processo de produção poderia ser melhorado?

### Resolução

Gráfico 54: Perfis médios da resposta de acordo com Máquina e Operador



Observando o gráfico X, pode-se melhorar o processo treinando os operadores da máquina 1, e o operador 7 da máquina 2, pois eles são os operadores que menos engarrafam bebidas, o que pode ser um treinamento inadequado.

- (e) O operador 4, avaliado na Máquina 1, tem pouca experiência comparado com os outros três operadores. Obtenha uma estimativa intervalar do seguinte contraste:

$$(\mu_{11} + \mu_{12} + \mu_{13})/3 - \mu_{14}$$

### Resolução

Uma Estimativa intervalar para o contraste  $(\mu_{11} + \mu_{12} + \mu_{13})/3 - \mu_{14}$  é

$$\sum_{j=1}^k c_j \bar{Y}_j \pm t_{\nu, \alpha/2} S_c \sqrt{\sum_{j=1}^k \frac{c_j^2}{n_j}}$$

onde  $S_c = \sqrt{QMRes}$  e  $\nu = n - k$ , assim nós temos:

$C = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, -1)$ ,  $\nu = 20 - 4$ ,  $\bar{Y} = (61.8, 67.8, 62.6, 52.6)$ , escolhendo um nível de confiança de 95% ficamos com um  $t_{16;0.95} = 2.12$ ,  $S_c = 5.28$  e  $\sqrt{\sum_{j=1}^k \frac{c_j^2}{n_j}} = 0.5164$ . Agora com todos os elementos calculados uma estimativa intervalar para  $(\mu_{11} + \mu_{12} + \mu_{13})/3 - \mu_{14}$  com 95% de confiança é:

$$(61.8 + 67.8 + 62.6)/3 - 52.6 \pm 2.12 * 5.28 * 0.5164 \Rightarrow 11.4666 \pm 5.78$$



## Exercício 6

Considere o experimento em que questionários de três diferentes cores foram distribuídos aleatoriamente em estacionamentos de 15 supermercados em duas semanas. O objetivo é investigar se há diferença na taxa de retorno de acordo com a cor. Os dados são apresentados a seguir. Considere que de acordo com as cores das marcas dos supermercados, houve restrição na atribuição das cores dos questionários a serem usados. Assim, realize a análise sob uma estrutura hierárquica de tratamentos (com cor e supermercados como fatores fixos). Considere semana como bloco.

Cor		Azul					Verde					Laranja				
Estacionamento		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Semana	1	28	26	31	27	35	34	29	25	31	29	31	25	27	29	28
	2	32	23	29	24	37	33	27	22	34	25	35	28	25	25	31

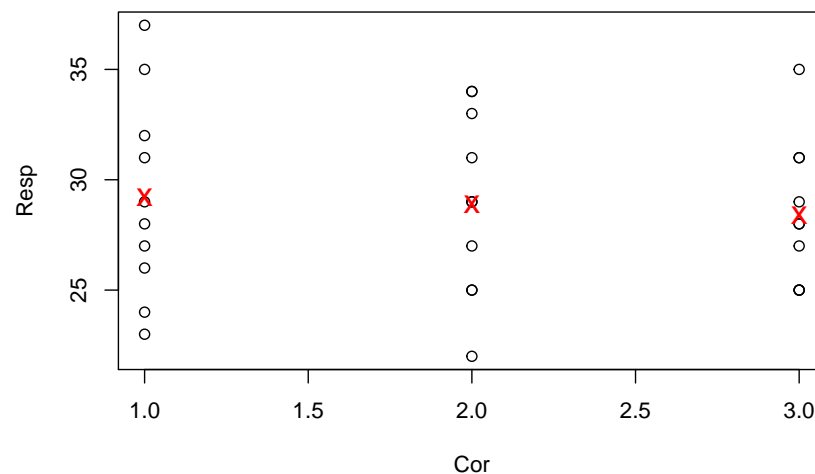
## Resolução

Foi criado um modelo do tipo:

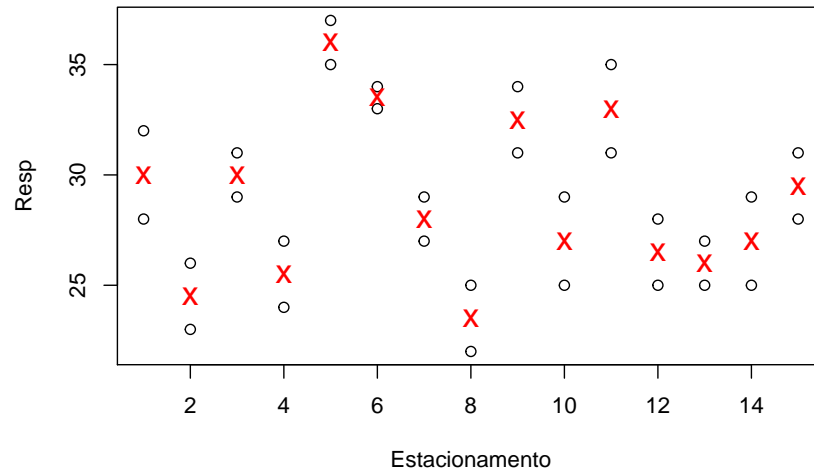
$$Y_{ijkl} = \mu + \tau_j + \beta_{i(j)} + \gamma_l + e_{ijkl}$$

em que  $\mu$  é a média de todos os valores, os tratamentos são as cores e hierárquicamente os estacionamentos e os blocos são as semanas 1 e 2. Realizado estatísticas descritivas para o modelo temos os seguintes gráficos:

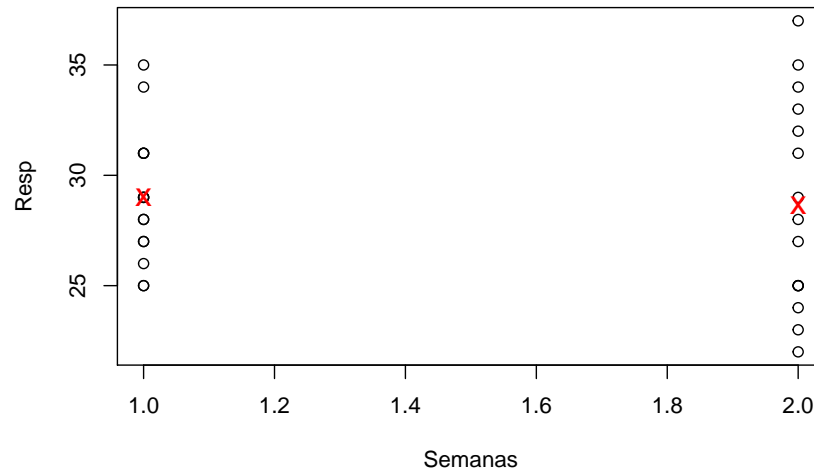
**Gráfico 55: Taxa de retorno de acordo com a Cor**



**Gráfico 56: Taxa de retorno de acordo com a Cor**



**Gráfico 57: Taxa de retorno de acordo com a Cor**



A partir desses gráficos podemos dizer que aparentemente não há diferença entre as cores, alguns estacionamentos receberam taxas muito maiores que outros e na primeira semana a variabilidade foi menor do que na segunda semana. Construímos então a tabela de ANOVA:

Table 13: Tabela de ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Cor	2	3.2666667	1.6333333	0.3430	0.7154373
Semanas	1	0.8333333	0.8333333	0.1750	0.6820520
Cor:Estacionamento	12	369.4000000	30.7833333	6.4645	0.0007569
Residuals	14	66.6666667	4.7619048	NA	NA

Realizamos ainda a análise de resíduos:

Gráfico 58: Resíduos vs Preditos

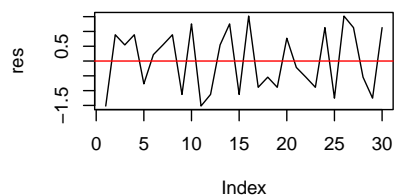


Gráfico 59: Resíduos vs Preditos

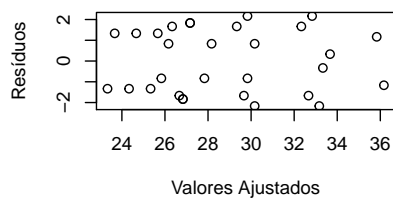


Gráfico 60: Histograma dos resíduos

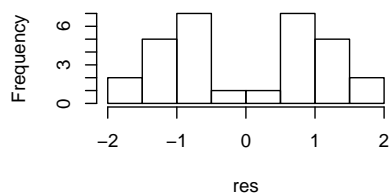


Gráfico 61: QQPlot Normal

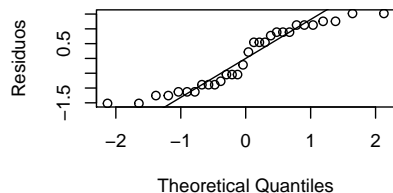


Gráfico 62: Resíduos vs B(A)

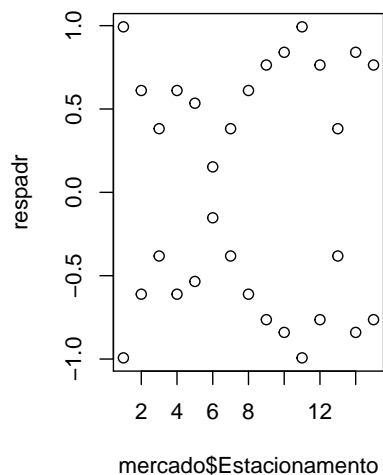
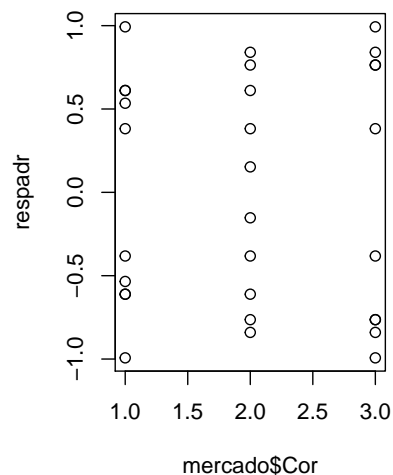


Gráfico 63: Resíduos vs A



Aparentemente não existe normalidade nos resíduos. As estimativas dos parâmetros são dadas por:

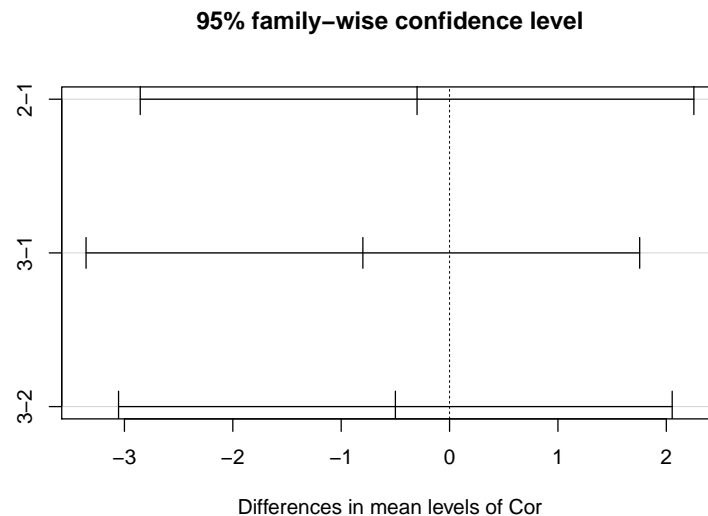
```
##      (Intercept)          Cor2          Cor3
##      3.016667e+01      -3.000000e+00      -5.000000e-01
##      Semanas2 Cor1:Estacionamento2 Cor2:Estacionamento2
##      -3.333333e-01      -5.500000e+00      NA
## Cor3:Estacionamento2 Cor1:Estacionamento3 Cor2:Estacionamento3
##      NA      1.994285e-15      NA
## Cor3:Estacionamento3 Cor1:Estacionamento4 Cor2:Estacionamento4
##      NA      -4.500000e+00      NA
## Cor3:Estacionamento4 Cor1:Estacionamento5 Cor2:Estacionamento5
##      NA      6.000000e+00      NA
## Cor3:Estacionamento5 Cor1:Estacionamento6 Cor2:Estacionamento6
##      NA      NA      6.500000e+00
```

```
## Cor3:Estacionamento6 Cor1:Estacionamento7 Cor2:Estacionamento7
## NA NA 1.000000e+00
## Cor3:Estacionamento7 Cor1:Estacionamento8 Cor2:Estacionamento8
## NA NA -3.500000e+00
## Cor3:Estacionamento8 Cor1:Estacionamento9 Cor2:Estacionamento9
## NA NA 5.500000e+00
## Cor3:Estacionamento9 Cor1:Estacionamento10 Cor2:Estacionamento10
## NA NA NA
## Cor3:Estacionamento10 Cor1:Estacionamento11 Cor2:Estacionamento11
## NA NA NA
## Cor3:Estacionamento11 Cor1:Estacionamento12 Cor2:Estacionamento12
## 3.500000e+00 NA NA
## Cor3:Estacionamento12 Cor1:Estacionamento13 Cor2:Estacionamento13
## -3.000000e+00 NA NA
## Cor3:Estacionamento13 Cor1:Estacionamento14 Cor2:Estacionamento14
## -3.500000e+00 NA NA
## Cor3:Estacionamento14 Cor1:Estacionamento15 Cor2:Estacionamento15
## -2.500000e+00 NA NA
## Cor3:Estacionamento15
## NA
```

Fazendo então um estudo do efeito principal do fator cor:

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = Resp ~ Cor/Estacionamento + Semanas)
##
## $Cor
##      diff      lwr      upr      p adj
## 2-1 -0.3 -2.854204 2.254204 0.9494277
## 3-1 -0.8 -3.354204 1.754204 0.6973386
## 3-2 -0.5 -3.054204 2.054204 0.8666151
```

Gráfico 64: Tukey efeito de tratamento



Concluimos então que não há diferença entre as cores utilizadas.

```

# Lista 4 - MAE0317

library(car)
library(MASS)

# Exercício 1

# item b

Pulse <- read.csv("Pulse.txt", header=T, sep= ";")
attach(Pulse)
n1 <- 35
n2 <- 57
Tratb <- factor(c(rep(1,n1),rep(2,n2)))
anovb <- data.frame(P2,Tratb)
boxplot(P2~Tratb,xlab="Correu")
title("Boxplot de P2 por Tratamento")

modb <- aov(P2~Tratb,data= anovb)
anovab <- anova(modb)
knitr::kable(caption = "Tabela de ANOVA", anovab)

s2 <- anovab$"Mean Sq"[2]
par(mfrow=c(2,2))
res <- rstudent(modb) # extraíndo resíduos
# Verificação de independência
plot(res, type="l")
title("Gráfico 2:Resíduos vs índices")
abline(h=0, col="red")

plot(modb$fit, modb$res, xlab="Valores Ajustados", ylab="Resíduos")
title("Gráfico 3:Resíduos vs Preditos")

hist(res, main=NULL)
title("Gráfico 4:Histograma dos resíduos")

qqnorm(res,ylab="Resíduos", main=NULL)
qqline(res)
title("Gráfico 5: QQPPlot Normal")

#Teste de normalidade de Shapiro-Wilk
shapiro.test(res)

# Teste de homogeneidade de variâncias

leveneTest(modb)
bartlett.test(P2~Tratb)

# item d

Dif <- P2 - P1

Tratd <- factor(c(rep(1,n1),rep(2,n2)))

```

```

anovd <- data.frame(Dif,Tratd)
boxplot(Dif~Tratd,xlab="Correu")
title("Boxplot de Dif=P2-P1 por Tratamento")

modd <- aov(Dif~Tratd,data= anovd)
anovad <- anova(modd)
knitr::kable(caption = "Tabela de ANOVA", anovad)

s2 <- anovad$"Mean Sq"[2]

par(mfrow=c(2,2))
res <- rstudent(modd) # extraindo resíduos
# Verificação de independência
plot(res, type="l") # res x índice das observações
title("Gráfico 7: Resíduos vs índices")
abline(h=0, col="red")

plot(modd$fit, modd$res, xlab="Valores Ajustados", ylab="Resíduos")
title("Gráfico 8:Resíduos vs Preditos")

hist(res, main=NULL)
title("Gráfico 9:Histograma dos resíduos")

qqnorm(res,ylab="Residuos", main=NULL)
qqline(res)
title("Gráfico 10: QQPlot Normal")

#Teste de normalidade de Shapiro-Wilk
shapiro.test(res)

# Teste de homogeneidade de variâncias
leveneTest(modd)

# item e

Medidas <- c(P1[1:35],rep(0,92),P1[36:92],P2[1:35],rep(0,92),P2[36:92])
Bloco <- factor(c(1:92,1:92,1:92,1:92))
correu_ou_ncorreu <- factor(c(rep(1,92),rep(2,92),rep(1,92),rep(2,92)))
inicial_ou_final <- factor(c(rep(1,184),rep(2,184)))
dadosd <- data.frame(Medidas,correu_ou_ncorreu,inicial_ou_final,Bloco)
boxplot(Medidas~correu_ou_ncorreu,xlab="Correu")
title("Boxplot das Medidas por Tratamento se Correu ou não")

boxplot(Medidas~inicial_ou_final,xlab="Tempo")
title("Boxplot das Medidas por Tratamento se é tempo inicial ou final")

mode <- aov(Medidas~correu_ou_ncorreu*inicial_ou_final+Bloco)
anovae <- anova(mode)
knitr::kable(caption = "Tabela de ANOVA", anovae)

s2 <- anovae$"Mean Sq"[5]

```

```

par(mfrow=c(2,2))
res <- rstudent(mode) # extraindo resíduos
# Verificação de independência
plot(res, type="l") # res x índice das observações
title("Gráfico 13: Resíduos vs índices")
abline(h=0, col="red")

plot(mode$fit, mode$res, xlab="Valores Ajustados", ylab="Resíduos")
title("Gráfico 14: Resíduos vs Preditos")

hist(res, main=NULL)
title("Gráfico 15: Histograma dos resíduos")

qqnorm(res, ylab="Resíduos", main=NULL)
qqline(res)
title("Gráfico 16: QQPPlot Normal")

# item f

Medidas <- c(P1,P2)
correu_ou_ncorreu <- factor(c(rep(1,35),rep(2,57),rep(1,35),rep(2,57)))
inicial_ou_final <- factor(c(rep(1,92),rep(2,92)))
Bloco <- factor(c(1:92,1:92))
dadosf <- data.frame(Medidas,correu_ou_ncorreu,inicial_ou_final,Bloco)
boxplot(Medidas~correu_ou_ncorreu,xlab="Correu")
title("Boxplot das Medidas por Tratamento se Correu ou não")

boxplot(Medidas~inicial_ou_final,xlab="Tempo")
title("Boxplot das Medidas por Tratamento se é tempo inicial ou final")

modf <- aov(Medidas~correu_ou_ncorreu/inicial_ou_final+Bloco)
anovaf <- anova(modf)
knitr::kable(caption = "Tabela de ANOVA", anovaf)

s2 <- anovaf$"Mean Sq"[4]
par(mfrow=c(2,2))
res <- rstudent(modf) # extraindo resíduos
# Verificação de independência
plot(res, type="l") # res x índice das observações
title("Gráfico 19: Resíduos vs índices")
abline(h=0, col="red")

plot(modf$fit, modf$res, xlab="Valores Ajustados", ylab="Resíduos")
title("Gráfico 20: Resíduos vs Preditos")

hist(res, main=NULL)
title("Gráfico 21: Histograma dos resíduos")

qqnorm(res, ylab="Resíduos", main=NULL)
qqline(res)
title("Gráfico 22: QQPPlot Normal")

par(mfrow=c(1,2))

```

```

plot.default(inicial_ou_final,res)
title("Gráfico 23: Resíduos vs B(A)")
plot.default(correu_ou_ncorreu,res)
title("Gráfico 24: Resíduos vs A")

#Teste de normalidade de Shapiro-Wilk
shapiro.test(res)

# Teste de homogeneidade de variâncias
leveneTest(Medidas~correu_ou_ncorreu/inicial_ou_final)

# item g

Medidas <- c(P1,P2)
correu_ou_ncorreu <- factor(c(rep(1,35),rep(2,57),rep(1,35),rep(2,57)))
inicial_ou_final <- factor(c(rep(1,92),rep(2,92)))
Bloco <- factor(c(1:92,1:92))
dadosg <- data.frame(Medidas,correu_ou_ncorreu,inicial_ou_final,Bloco)

boxcox(Medidas ~ correu_ou_ncorreu/inicial_ou_final+Bloco, data=dadosf, plotit=T)

Medidas <- Medidas^-1
dadosg <- data.frame(Medidas,correu_ou_ncorreu,inicial_ou_final,Bloco)
modg <- aov(Medidas~correu_ou_ncorreu/inicial_ou_final+Bloco)
anovag <- anova(modg)
knitr::kable(caption = "Tabela de ANOVA", anovag)

s2 <- anovag$"Mean Sq"[4]

par(mfrow=c(2,2))
res <- rstudent(modg) # extraíndo resíduos
# Verificação de independência
plot(res, type="l") # res x índice das observações
title("Gráfico 24: Resíduos vs índices")
abline(h=0, col="red")

plot(modg$fit, modg$res, xlab="Valores Ajustados", ylab="Resíduos")
title("Gráfico 25: Resíduos vs Preditos")

hist(res, main=NULL)
title("Gráfico 26: Histograma dos resíduos")

qqnorm(res,ylab="Residuos", main=NULL)
qqline(res)
title("Gráfico 27: QQPLOT Normal")

par(mfrow=c(1,2))
plot.default(inicial_ou_final,res)
title("Gráfico 28: Resíduos vs B(A)")
plot.default(correu_ou_ncorreu,res)
title("Gráfico 29: Resíduos vs A")

```



```

#Teste de normalidade de Shapiro-Wilk
shapiro.test(res)

# Teste de homogeneidade de variâncias
leveneTest(Medidas~correu_ou_ncorreu/inicial_ou_final)

# Exercício 3

# item a

resp <- c(0.02,0.15,0.45,0.18,0.27,0.24,-0.01,0.58,0.11,0.35,0.14,-0.03,0.48,0.04,0.18,0.22)
Bloco <- factor(c(rep(1,4),rep(2,4),rep(3,4),rep(4,4)))
Var <- factor(c(1,2,3,4,2,3,1,4,3,4,2,1,4,1,3,2))
dat <- data.frame(resp,Var,Bloco)

#ANOVA do DABC
fit1 <- aov(resp ~ Var + Bloco)
as <- anova(fit1)
knitr::kable(caption = "Tabela de ANOVA", as)
res <- rstudent(fit1)

par(mfrow=c(2,2))
plot(fit1$fitted.values,res)
title("Gráfico 30: Resíduos vs índices")
plot.default(dat$Bloco,fit1$residuals)
title("Gráfico 31: Resíduos vs Blocos")
plot.default(dat$Var,fit1$residuals)
title("Gráfico 32: Resíduos vs Tratamentos")
qqnorm(fit1$residuals,ylab="Resíduos", main=NULL)
qqline(fit1$residuals)
title("Gráfico 33: QQPlot Normal")

# item b

interaction.plot(Bloco, Var, resp, main="Gráfico 34: Efeito de Interação dos Fatores Var e Bloco")

# item e

tr <- c(0, fit1$coeff[2:4])
tr <- rep(tr, 4)
bl <- c(0, fit1$coeff[5:7])
bl <- rep(bl, each=4)
trbl <- tr * bl #novo fator que mede a interação
datt <- cbind(dat,tr,bl, trbl)

fitna <- aov(resp ~ Var + Bloco + trbl)
d <- anova(fitna) #Interprete. O que está sendo testado?
knitr::kable(caption = "Tabela de ANOVA", d)

#Estudo do efeito de Tratamento
fit1.tk <- TukeyHSD(fit1, "Var")
fit1.tk

```

```

plot(fit1.tk)

# Exercício 4

# Exp1
Resp1 <- matrix(c(2 , 3 , 4 , 6,
                  6 , 8 , 8 , 12,
                  13, 10, 15, 20),12,1)

Bloco <- factor(c(rep(1:4,each=3)))
Trat <- factor(c(rep(1:3,4)))
tab1 <- cbind(Bloco,Trat,Resp1)
tab1 <- data.frame(tab1)

#ANOVA do DABC
fit0 <- aov(Resp1 ~ Trat + Bloco)
an <- anova(fit0)
knitr::kable(caption = "Tabela de ANOVA", an)

interaction.plot(Bloco, Trat, Resp1, main="Efeito de Interação dos Fatores Tratamento e Bloco")

par(mfrow=c(2,2))
plot(fit0$fitted.values,rstudent(fit0), ylab="Resíduos studentizados", xlab="Valores Preditos")
title("Gráfico 37: Resíduos vs Preditos")
plot.default(tab1$Bloco,rstudent(fit0), ylab="Resíduos studentizados", xlab="Blocos")
title("Gráfico 38: Resíduos vs Blocos")
plot.default(tab1$Trat,rstudent(fit0), ylab="Resíduos studentizados", xlab="Tratamento")
title("Gráfico 39: Resíduos vs Tratamento")
plot(rstudent(fit0), type="l", ylab="Resíduos studentizados", xlab="Índices")
title("Gráfico 40: Resíduos vs índice")
abline(h=0, col="red")
par(mfrow=c(1,1))
qqnorm(rstudent(fit0),ylab="Resíduos studentizados", main=NULL)
qqline(rstudent(fit0))
title("Gráfico 41: QQPlot Normal")

tr <- c(0, fit0$coeff[2:3])
tr <- rep(tr, 4)
bl <- c(0, fit0$coeff[4:6])
bl <- rep(bl, each=3)
trbl <- tr * bl # Novo fator que mede a interação
datt <- cbind(tab1,tr,bl, trbl)

fitna <- aov(Resp1 ~ Trat + Bloco + trbl)
anv <- anova(fitna)
knitr::kable(caption = "Tabela de ANOVA", anv)

fit0.tka <- TukeyHSD(fit0, "Trat")
fit0.tka
plot(fit0.tka)

# Exp2
Resp2 <- matrix(c(2 , 15, 16, 6,

```

```

12, 13, 8 , 6,
7 , 10, 3 , 4),12,1)

tab2 <- cbind(Bloco,Trat,Resp2)
tab2 <- data.frame(tab2)

#ANOVA do DABC
fit1 <- aov(Resp2 ~ Trat + Bloco)
an <- anova(fit1)
knitr::kable(caption = "Tabela de ANOVA", an)

interaction.plot(Bloco, Trat, Resp2, main="Efeito de Interação dos Fatores Tratamento e Bloco")

par(mfrow=c(2,2))
plot(fit1$fitted.values,rstudent(fit1), ylab="Resíduos studentizados", xlab="Valores Preditos")
title("Gráfico 44: Resíduos vs Preditos")
plot.default(tab2$Bloco,rstudent(fit1), ylab="Resíduos studentizados", xlab="Blocos")
title("Gráfico 45: Resíduos vs Blocos")
plot.default(tab2$Trat,rstudent(fit1), ylab="Resíduos studentizados", xlab="Tratamento")
title("Gráfico 46: Resíduos vs Tratamento")
plot(rstudent(fit1), type="l", ylab="Resíduos studentizados", xlab="Índices")
title("Gráfico 47:Resíduos vs índice")
abline(h=0, col="red")
par(mfrow=c(1,1))
qqnorm(rstudent(fit1),ylab="Resíduos studentizados", main=NULL)
qqline(rstudent(fit1))
title("Gráfico 48: QQPlot Normal")

tr <- c(0, fit1$coeff[2:3])
tr <- rep(tr, 4)
bl <- c(0, fit1$coeff[4:6])
bl <- rep(bl, each=3)
trbl2 <- tr * bl #novo fator que mede a interação
datt <-cbind(tab2,tr,bl, trbl2)

fitna <- aov(Resp2 ~ Trat + Bloco + trbl2)
anvo <- anova(fitna)
knitr::kable(caption = "Tabela de ANOVA", anvo)

# Exercício 5

# item a

Resp <- c(65,58,63,57,66,68,62,75,64,70,56,
65,58,70,64,45,56,54,48,60,74,81,
76,80,68,69,76,80,78,73,52,56,62,
58,51,73,78,83,75,76,69,83,74,78,
80,63,70,72,68,75,81,72,73,76,70,
67,79,73,77,71)

dados <- data.frame(fa=factor(rep(1:3,each=20)), fb=factor(rep(1:12,each=5)), resp=Resp)
attach(dados)
fit1 <- aov(Resp ~ fa/fb, data=dados)
a <- anova(fit1)

```

```

knitr::kable(caption = "Tabela de ANOVA", a)

par(mfrow=c(2,2))
plot(fit1$fitted.values,rstudent(fit1), ylab="Resíduos studentizados", xlab="Valores Preditos")
title("Gráfico 49: Resíduos vs Preditos")
plot.default(dados$fb,rstudent(fit1))
title("Gráfico 50: Resíduos vs Operadores(Máquinas)")
plot.default(dados$fa,rstudent(fit1))
title("Gráfico 51: Resíduos vs Máquinas")
plot(rstudent(fit1), type="l", ylab="Resíduos studentizados", xlab="Índices")
title("Gráfico 52:Resíduos vs índices")
abline(h=0, col="red")

qqnorm(rstudent(fit1),ylab="Resíduos studentizados", main=NULL)
qqline(rstudent(fit1))
title("Gráfico 53: QQPlot Normal")

# item c

s2 <- sum(resid(fit1)^2)/fit1$df.res
dt <- qtkey(0.95, 2, 48) * sqrt(s2/5)
fit1.m <- tapply(Resp, list(fa,fb), mean)
# Para maquina 1
md1 <- fit1.m[1,1]-fit1.m[1,2]
data.frame("1-2",dif=md1, li=md1-dt, ls=md1+dt, sig=ifelse(abs(md1)>dt,"*", "ns"))
md2 <- fit1.m[1,1]-fit1.m[1,3]
data.frame("1-3",dif=md2, li=md2-dt, ls=md2+dt, sig=ifelse(abs(md2)>dt,"*", "ns"))
md3 <- fit1.m[1,1]-fit1.m[1,4]
data.frame("1-3",dif=md3, li=md3-dt, ls=md3+dt, sig=ifelse(abs(md3)>dt,"*", "ns"))
md4 <- fit1.m[1,2]-fit1.m[1,3]
data.frame("2-3",dif=md4, li=md4-dt, ls=md4+dt, sig=ifelse(abs(md4)>dt,"*", "ns"))
md5 <- fit1.m[1,2]-fit1.m[1,4]
data.frame("2-4",dif=md5, li=md5-dt, ls=md5+dt, sig=ifelse(abs(md5)>dt,"*", "ns"))
md6 <- fit1.m[1,3]-fit1.m[1,4]
data.frame("3-4",dif=md6, li=md6-dt, ls=md6+dt, sig=ifelse(abs(md6)>dt,"*", "ns"))

# Para maquina 2
md7 <- fit1.m[2,5]-fit1.m[2,6]
data.frame("5-6",dif=md7, li=md7-dt, ls=md7+dt, sig=ifelse(abs(md7)>dt,"*", "ns"))
md8 <- fit1.m[2,5]-fit1.m[2,7]
data.frame("5-7",dif=md8, li=md8-dt, ls=md8+dt, sig=ifelse(abs(md8)>dt,"*", "ns"))
md9 <- fit1.m[2,5]-fit1.m[2,8]
data.frame("5-8",dif=md9, li=md9-dt, ls=md9+dt, sig=ifelse(abs(md9)>dt,"*", "ns"))
md10 <- fit1.m[2,6]-fit1.m[2,7]
data.frame("6-7",dif=md10, li=md10-dt, ls=md10+dt, sig=ifelse(abs(md10)>dt,"*", "ns"))
md11 <- fit1.m[2,6]-fit1.m[2,8]
data.frame("6-8",dif=md11, li=md11-dt, ls=md11+dt, sig=ifelse(abs(md11)>dt,"*", "ns"))
md12 <- fit1.m[2,7]-fit1.m[2,8]
data.frame("7-8",dif=md12, li=md12-dt, ls=md12+dt, sig=ifelse(abs(md12)>dt,"*", "ns"))

# Para maquina 3
md13 <- fit1.m[3,9]-fit1.m[3,10]

```

```

data.frame("9-10",dif=md13, li=md13-dt, ls=md13+dt, sig=ifelse(abs(md13)>dt,"*", "ns"))
md14 <- fit1.m[3,9]-fit1.m[3,11]
data.frame("9-11",dif=md14, li=md14-dt, ls=md14+dt, sig=ifelse(abs(md14)>dt,"*", "ns"))
md15 <- fit1.m[3,9]-fit1.m[3,12]
data.frame("9-12",dif=md15, li=md15-dt, ls=md15+dt, sig=ifelse(abs(md15)>dt,"*", "ns"))
md16 <- fit1.m[3,10]-fit1.m[3,11]
data.frame("10-11",dif=md16, li=md16-dt, ls=md16+dt, sig=ifelse(abs(md16)>dt,"*", "ns"))
md17 <- fit1.m[3,10]-fit1.m[3,12]
data.frame("10-12",dif=md17, li=md17-dt, ls=md17+dt, sig=ifelse(abs(md17)>dt,"*", "ns"))
md18 <- fit1.m[3,11]-fit1.m[3,12]
data.frame("11-12",dif=md18, li=md18-dt, ls=md18+dt, sig=ifelse(abs(md18)>dt,"*", "ns"))

# item d

interaction.plot(fb, fa, Resp, main="Perfis médios de acordo com Máquina e Operador", xlab="Operadores")

# Exercício 6

# item a

Cor <- factor(c(rep(1,5),rep(2,5),rep(3,5),rep(1,5),rep(2,5),rep(3,5)))
Estacionamento <- factor(c(seq(1:15),seq(1:15)))
Semanas <- factor(c(rep(1,15),rep(2,15)))
Resp <- c(28, 26, 31, 27, 35, 34, 29, 25, 31, 29, 31, 25, 27, 29, 28, 32, 23, 29, 24, 37, 33, 27, 22, 30)

mercado <- data.frame(Cor,Estacionamento,Semanas,Resp)

mcor <- tapply(Resp,Cor,mean)
plot.default(Cor, Resp,main="Gráfico 55: Taxa de retorno de acordo com a Cor")
points(mcor, pch="x", col=2, cex=1.5)

mestacionamento <- tapply(Resp,Estacionamento,mean)
plot.default(Estacionamento, Resp,main="Gráfico 56: Taxa de retorno de acordo com a Cor")
points(mestacionamento, pch="x", col=2, cex=1.5)

msemana <- tapply(Resp,Semanas,mean)
plot.default(Semanas, Resp,main="Gráfico 57: Taxa de retorno de acordo com a Cor")
points(msemana, pch="x", col=2, cex=1.5)

mod6 <- aov(Resp ~ Cor/Estacionamento+Semanas)
anova6 <- anova(mod6)
knitr::kable(caption = "Tabela de ANOVA", anova6)

par(mfrow=c(2,2))
s2 <- anova6$"Mean Sq"[4]

par(mfrow=c(2,2))
res <- rstudent(mod6) # extraindo resíduos
# Verificação de independência
plot(res, type="l") # res x índice das observações
title("Gráfico 58: Resíduos vs Preditos")
abline(h=0, col="red")

```

```

plot(mod6$fit, mod6$res, xlab="Valores Ajustados", ylab="Resíduos")
title("Gráfico 59: Resíduos vs Preditos")

hist(res, main=NULL)
title("Gráfico 60: Histograma dos resíduos")

qqnorm(res, ylab="Resíduos", main=NULL)
qqline(res)
title("Gráfico 61: QQPPlot Normal")

respadr <- (mod6$residuals/sqrt(anova6$"Mean Sq"[4]))
par(mfrow=c(1,2))
plot.default(mercado$Estacionamento, respadr)
title("Gráfico 62: Resíduos vs B(A)")
plot.default(mercado$Cor, respadr)
title("Gráfico 63: Resíduos vs A")

mod6$coeff

mod6.tka <- TukeyHSD(mod6, "Cor")
mod6.tka

plot(mod6.tka)

```