

Análise de Dados DIC - Dados sobre gordura

Cleanderson Fidelis

O problema das rosquinhas

Durante o cozimento, as rosquinhas absorvem gordura em diferentes quantidades. Este estudo busca determinar se a **quantidade de gordura absorvida** pelas rosquinhas depende do tipo de gordura utilizada no cozimento. Foram avaliados:

- **4 tipos de gordura** (TG_1, TG_2, TG_3, TG_4).
- **6 lotes** de rosquinhas por tipo de gordura.
- Cada lote é composto por 24 rosquinhas.

Os dados são expressos em gramas de gordura absorvida por lote, subtraindo-se 100 g para simplificação. A variável **tipo de gordura** é o único fator de classificação neste experimento.

Dados do Estudo

Dados de Gordura Absorvida

Tipo de gordura	Lote 1	Lote 2	Lote 3	Lote 4	Lote 5	Lote 6	Total	Média
TG1	64	72	68	77	56	95	432	72
TG2	78	91	97	82	85	77	510	85
TG3	75	93	78	71	63	76	456	76
TG4	55	66	49	64	70	68	372	62
Geral	-	-	-	-	-	-	1770	73.75

Leitura dos dados no R

```
# Criando o data.frame com os dados fornecidos
dados_rosquinhas <- data.frame(
  Tipo_de_gordura = rep(c("TG1", "TG2", "TG3", "TG4"), each = 6), # Tipos de gordura
  Lote = rep(1:6, times = 4), # Lotes de 1 a 6 para cada tipo
  Gordura_absorvida = c(
    64, 72, 68, 77, 56, 95, # Gordura absorvida por TG1
    78, 91, 97, 82, 85, 77, # Gordura absorvida por TG2
    75, 93, 78, 71, 63, 76, # Gordura absorvida por TG3
    55, 66, 49, 64, 70, 68 # Gordura absorvida por TG4
  )
)

# Visualizar os dados
print(dados_rosquinhas)
```

```
##      Tipo_de_gordura Lote Gordura_absorvida
## 1      TG1          1              64
```

```
## 2      TG1      2      72
## 3      TG1      3      68
## 4      TG1      4      77
## 5      TG1      5      56
## 6      TG1      6      95
## 7      TG2      1      78
## 8      TG2      2      91
## 9      TG2      3      97
## 10     TG2      4      82
## 11     TG2      5      85
## 12     TG2      6      77
## 13     TG3      1      75
## 14     TG3      2      93
## 15     TG3      3      78
## 16     TG3      4      71
## 17     TG3      5      63
## 18     TG3      6      76
## 19     TG4      1      55
## 20     TG4      2      66
## 21     TG4      3      49
## 22     TG4      4      64
## 23     TG4      5      70
## 24     TG4      6      68
```

```
kableExtra::kable(dados_rosquinhas)
```

Tipo_de_gordura	Lote	Gordura_absorvida
TG1	1	64
TG1	2	72
TG1	3	68
TG1	4	77
TG1	5	56
TG1	6	95
TG2	1	78
TG2	2	91
TG2	3	97
TG2	4	82
TG2	5	85
TG2	6	77
TG3	1	75
TG3	2	93
TG3	3	78
TG3	4	71
TG3	5	63
TG3	6	76
TG4	1	55
TG4	2	66
TG4	3	49
TG4	4	64
TG4	5	70
TG4	6	68

Estatística Descritiva

A análise descritiva indica que o tipo de gordura TG2 apresentou a maior média de absorção (85 g) e a menor variabilidade (7,77 g), destacando-se pela consistência no processo. Em contrapartida, o tipo TG4 apresentou a menor média de absorção (62 g) e variabilidade moderada (8,22 g). O tipo TG1 mostrou a maior variabilidade (13,34 g), sugerindo maior heterogeneidade entre os lotes. Esses resultados preliminares apontam diferenças claras entre os tipos de gordura, sendo necessária uma análise inferencial, como ANOVA, para confirmar se essas diferenças são estatisticamente significativas.

```
# Estatísticas descritivas: Resumo geral e por tratamento
```

```
summary(dados_rosquinhas)
```

```
## Tipo_de_gordura      Lote      Gordura_absorvida
## Length:24           Min.      :1.0      Min.      :49.00
## Class :character     1st Qu.:2.0      1st Qu.:65.50
## Mode  :character     Median :3.5      Median :73.50
##                               Mean   :3.5      Mean   :73.75
##                               3rd Qu.:5.0      3rd Qu.:79.00
##                               Max.    :6.0      Max.    :97.00
```

```
aggregate(Gordura_absorvida ~ Tipo_de_gordura, data = dados_rosquinhas, FUN = function(x) c(mean = mean
```

```
## Tipo_de_gordura Gordura_absorvida.mean Gordura_absorvida.sd
## 1              TG1              72.000000          13.341664
## 2              TG2              85.000000           7.771744
## 3              TG3              76.000000           9.879271
## 4              TG4              62.000000           8.221922
```

Análises gráfica

Perfil

Os gráficos mostram a variação da gordura absorvida por lote para cada tipo de gordura (TG1, TG2, TG3, TG4). Observa-se que o tipo TG1 apresenta uma alta variabilidade entre os lotes, com um aumento significativo no Lote 6, sugerindo inconsistências nos resultados. O tipo TG2 mantém valores mais elevados e consistentes ao longo dos lotes, reforçando sua estabilidade. O tipo TG3 apresenta maior absorção no Lote 2, seguido de uma redução nos lotes subsequentes, indicando maior variabilidade em relação a TG2. Já o tipo TG4 possui a menor média de gordura absorvida, com oscilações mais discretas entre os lotes. Esses padrões indicam diferenças marcantes entre os tipos de gordura, destacando a necessidade de análises adicionais para compreender as causas dessas variações.

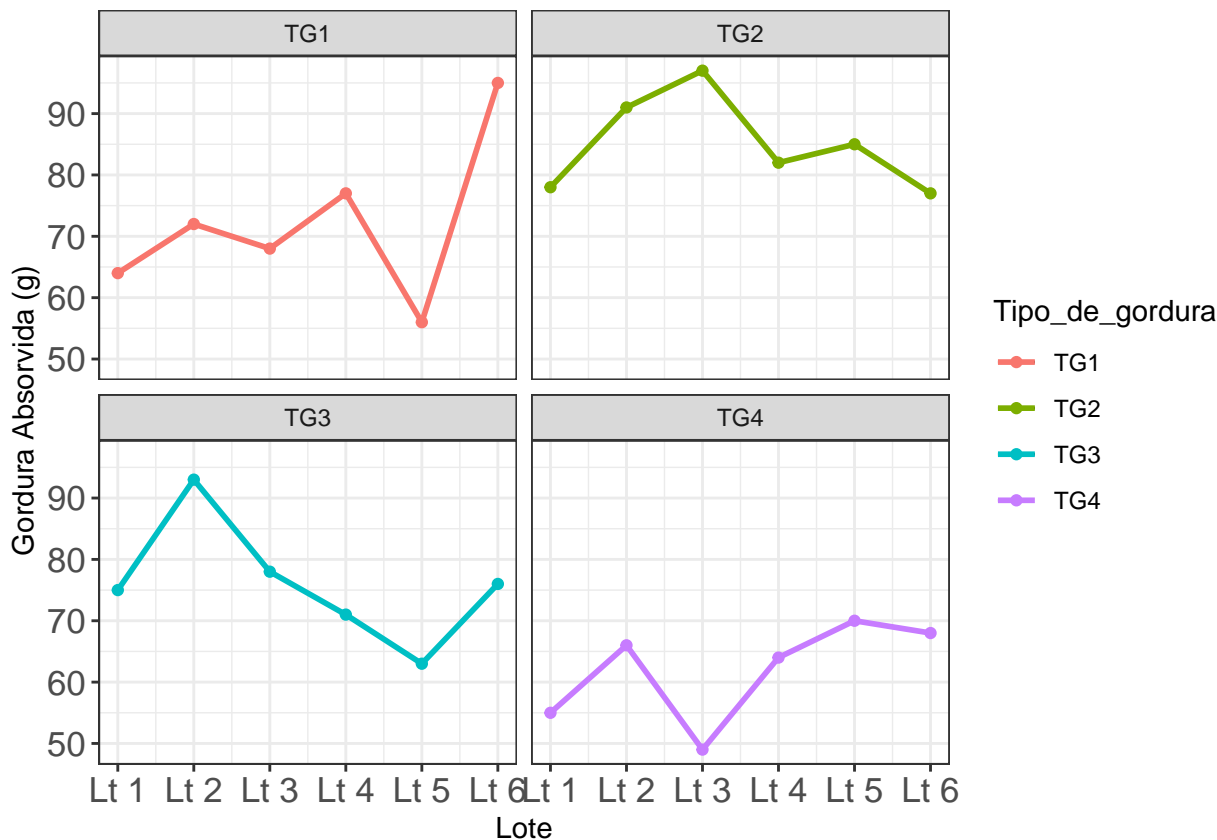
```
# Criar o gráfico para os dados de gordura absorvida
```

```
ggplot(dados_rosquinhas, aes(x = Lote, y = Gordura_absorvida, group = Tipo_de_gordura)) +
  geom_line(aes(linetype = Tipo_de_gordura),
            size = 0.5) + # Linhas conectando os pontos por tipo de gordura
  geom_point(aes(color = Tipo_de_gordura),
             size = 1.5) + # Pontos para cada valor de gordura absorvida
  scale_x_continuous(breaks = 1:6,
                     labels = paste("Lt", 1:6)) + # Eixo X representando os lotes
  stat_summary(fun = mean, geom = "line", lwd = 1.0,
               aes(group = Tipo_de_gordura,
                   color = Tipo_de_gordura)) + # Média por tipo

  xlab("Lote") +
  ylab("Gordura Absorvida (g)") +
  theme_bw() + # Tema com fundo branco
  theme(axis.text.x = element_text(size = 13)) + # Tamanho do texto no eixo X
  theme(axis.text.y = element_text(size = 13)) + # Tamanho do texto no eixo Y
```

```
facet_wrap(~Tipo_de_gordura) # Dividir o gráfico por tipo de gordura
```

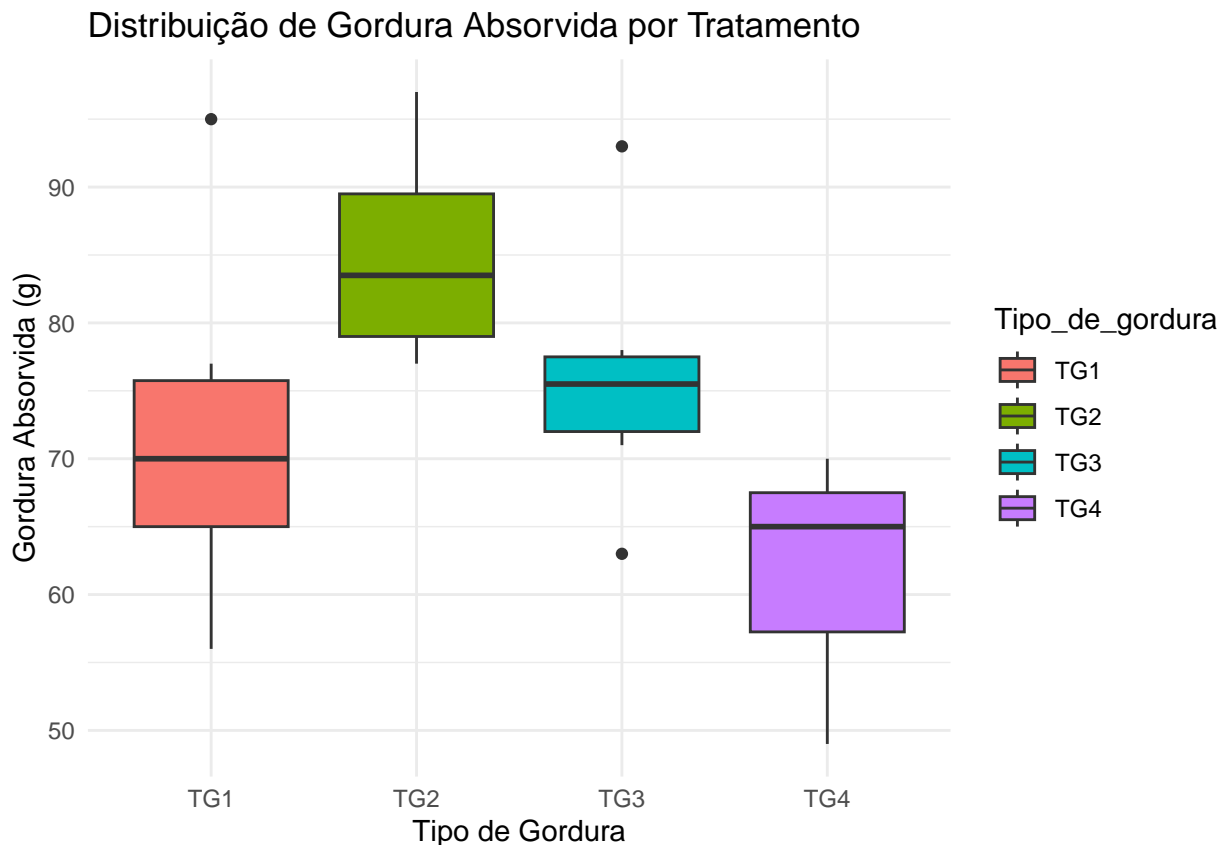
```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```



Box-plot

Os boxplots mostram a distribuição da gordura absorvida por lote para cada tipo de gordura (TG1, TG2, TG3, TG4). Observa-se que TG2 apresenta a maior mediana e uma distribuição relativamente concentrada, indicando maior absorção e menor variabilidade. TG1 possui uma mediana inferior a TG2, com maior amplitude e presença de valores extremos, sugerindo maior variabilidade entre os lotes. TG3 exibe uma mediana intermediária com menor dispersão, mas também apresenta valores discrepantes. Já TG4 apresenta a menor mediana, com uma amplitude relativamente ampla, refletindo menor absorção média e variabilidade moderada. Esses resultados reforçam as diferenças de desempenho entre os tipos de gordura, destacando TG2 como o mais consistente em termos de absorção.

```
# Gráfico de boxplot para visualizar a distribuição da gordura absorvida por tratamento
ggplot(dados_rosquinhas, aes(x = Tipo_de_gordura, y = Gordura_absorvida,
                             fill = Tipo_de_gordura)) +
  geom_boxplot() +
  labs(title = "Distribuição de Gordura Absorvida por Tratamento",
       x = "Tipo de Gordura",
       y = "Gordura Absorvida (g)") +
  theme_minimal()
```



Anova

- O fator “**Tratamento**” apresentou um valor $p = 0,0069$, indicando que, ao nível de significância de 5%, existe uma diferença estatisticamente significativa entre as médias dos tratamentos em relação à gordura absorvida.
- O F calculado para os tratamentos ($F_c = 5,4063$) é maior que o valor crítico esperado para essa configuração, reforçando a rejeição da hipótese nula de que todas as médias são iguais.
- O coeficiente de variação ($CV = 13,62\%$) indica uma variabilidade moderada nos dados em relação às médias observadas, o que é aceitável para experimentos desse tipo.
- O teste de Shapiro-Wilk resultou em um valor $p = 0,7205$, indicando que os resíduos seguem uma distribuição normal ao nível de significância de 5%.
- Isso cumpre uma das premissas fundamentais para a validade da ANOVA.
- O teste de Bartlett apresentou um valor $p = 0,6258$, sugerindo que as variâncias dos tratamentos são homogêneas ao nível de significância de 5%.
- Essa é outra premissa atendida para a aplicação da ANOVA.

Teste de Comparação Múltipla (Tukey)

O teste de Tukey agrupou os tratamentos em diferentes níveis de significância:

- **Grupo “a”**: TG2 apresentou a maior média (85), sendo significativamente superior aos demais.
- **Grupo “ab”**: TG3 (76) e TG1 (72) não diferem significativamente entre si, mas possuem médias inferiores ao TG2.
- **Grupo “b”**: TG4 apresentou a menor média (62), sendo significativamente inferior ao TG2 e ligeiramente inferior aos demais.

Os resultados da ANOVA e do teste de Tukey indicam que o tipo de gordura influencia significativamente a quantidade de gordura absorvida:

- O tipo TG2 mostrou o melhor desempenho, com a maior absorção e diferença estatística clara em relação ao TG4.
- Os tipos TG1 e TG3 apresentaram desempenhos intermediários, sem diferenças significativas entre si.
- A normalidade dos resíduos e a homogeneidade das variâncias confirmam que as premissas para a ANOVA foram atendidas, validando os resultados obtidos.
- Investigar os fatores específicos que levam ao melhor desempenho de TG2, como suas propriedades químicas.
- Considerar testes adicionais ou ajustes no modelo caso outras variáveis tenham influência nos resultados.

```
# Realizar a ANOVA com o pacote ExpDes.pt
# Verifica se há diferença significativa entre os tratamentos
dic(
  trat = dados_rosquinhas$Tipo_de_gordura,
  resp = dados_rosquinhas$Gordura_absorvida,
  quali = TRUE, # Tratamento é qualitativo
  mcomp = "tukey", # Teste de comparação múltipla: Tukey
  sigT = 0.05, # Nível de significância para Tukey
  sigF = 0.05 # Nível de significância para ANOVA
)
```

```
## -----
## Quadro da analise de variancia
## -----
##          GL      SQ      QM      Fc      Pr>Fc
## Tratamento  3 1636.5 545.5 5.4063 0.0068759
## Residuo      20 2018.0 100.9
## Total        23 3654.5
## -----
## CV = 13.62 %
##
## -----
## Teste de normalidade dos residuos ( Shapiro-Wilk )
## Valor-p: 0.7205129
## De acordo com o teste de Shapiro-Wilk a 5% de significancia, os residuos podem ser considerados norm
## -----
##
## -----
## Teste de homogeneidade de variancia
## valor-p: 0.6257791
## De acordo com o teste de bartlett a 5% de significancia, as variancias podem ser consideradas homoge
## -----
##
## Teste de Tukey
## -----
## Grupos Tratamentos Medias
## a      TG2      85
## ab     TG3      76
## ab     TG1      72
## b      TG4      62
## -----
```

Verificação dos pressupostos

```

# Verificar os pressupostos do modelo ANOVA
modelo <- aov(Gordura_absorvida ~ Tipo_de_gordura, data = dados_rosquinhas) # Ajustar modelo ANOVA

# Teste de normalidade dos resíduos (Shapiro-Wilk)
shapiro.test(residuals(modelo)) # p > 0.05 indica normalidade

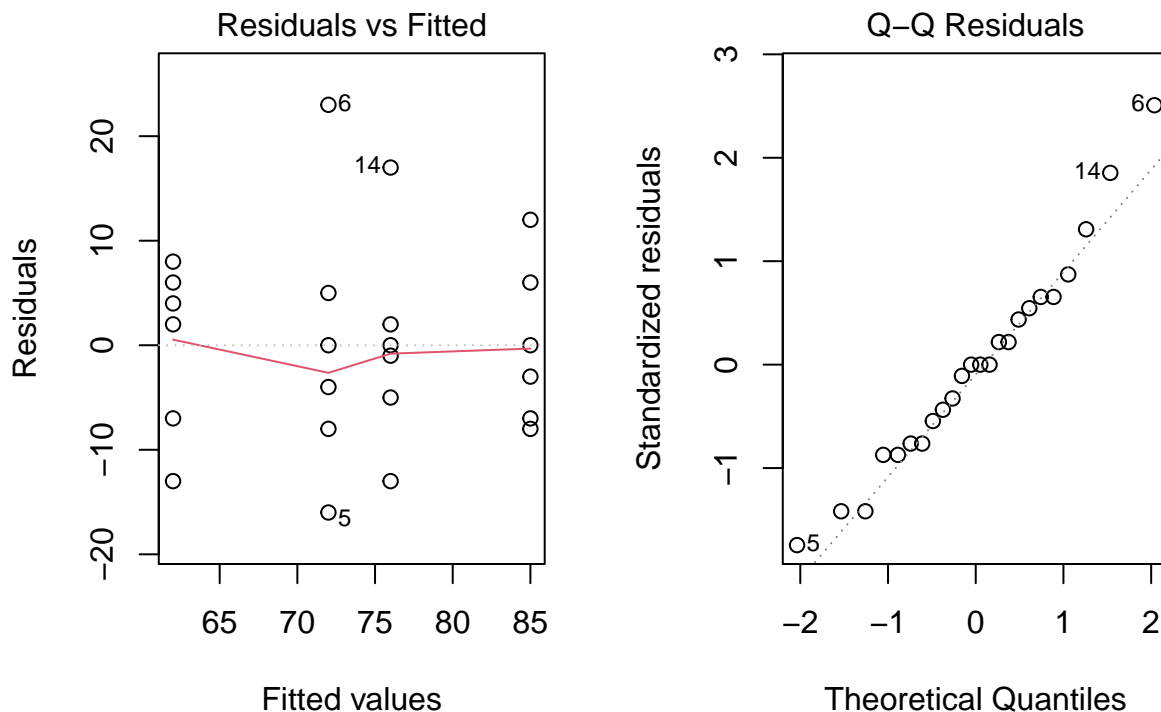
##
## Shapiro-Wilk normality test
##
## data: residuals(modelo)
## W = 0.97217, p-value = 0.7205

# Teste de homogeneidade de variâncias (Bartlett)
bartlett.test(Gordura_absorvida ~ Tipo_de_gordura,
              data = dados_rosquinhas) # p > 0.05 indica homogeneidade

##
## Bartlett test of homogeneity of variances
##
## data: Gordura_absorvida by Tipo_de_gordura
## Bartlett's K-squared = 1.7504, df = 3, p-value = 0.6258

# Plotar gráficos de resíduos
par(mfrow = c(1, 2)) # Configurar para dois gráficos lado a lado
plot(modelo, which = 1) # Resíduos vs Ajustados
plot(modelo, which = 2) # QQ plot dos resíduos

```



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

Utilizando o teste Scott-Knot

O teste de Scott-Knott é amplamente utilizado para formar grupos homogêneos de médias em experimentos com diferentes tratamentos. Nos resultados apresentados, o teste identificou dois grupos distintos, baseados

na absorção de gordura pelos tratamentos.

```
# Teste post-hoc: Scott-Knott (usando ExpDes.pt)
dic(
  trat = dados_rosquinhas$Tipo_de_gordura,
  resp = dados_rosquinhas$Gordura_absorvida,
  quali = TRUE,
  mcomp = "sk", # Teste Scott-Knott
  sigT = 0.05,
  sigF = 0.05
)
```

```
## -----
## Quadro da analise de variancia
## -----
##          GL      SQ      QM      Fc      Pr>Fc
## Tratamento  3 1636.5 545.5 5.4063 0.0068759
## Residuo      20 2018.0 100.9
## Total        23 3654.5
## -----
## CV = 13.62 %
##
## -----
## Teste de normalidade dos residuos ( Shapiro-Wilk )
## Valor-p:  0.7205129
## De acordo com o teste de Shapiro-Wilk a 5% de significancia, os residuos podem ser considerados norm
## -----
##
## -----
## Teste de homogeneidade de variancia
## valor-p:  0.6257791
## De acordo com o teste de bartlett a 5% de significancia, as variancias podem ser consideradas homogen
## -----
##
## Teste de Scott-Knott
## -----
##   Grupos Tratamentos Medias
## 1      a          TG2      85
## 2      a          TG3      76
## 3      a          TG1      72
## 4      b          TG4      62
## -----
```

O Grupo “a” inclui os tratamentos TG2, TG3 e TG1, que não apresentaram diferenças significativas entre suas médias ao nível de significância considerado. O tratamento TG2 teve a maior média de absorção (85), seguido por TG3 (76) e TG1 (72). Embora TG2 tenha a maior média, todos os três tratamentos demonstraram comportamentos estatisticamente similares em termos de absorção de gordura, justificando seu agrupamento no mesmo grupo.

Por outro lado, o Grupo “b” incluiu exclusivamente o tratamento TG4, que apresentou a menor média de absorção (62). Essa separação reflete uma diferença estatisticamente significativa entre TG4 e os demais tratamentos, indicando que TG4 possui características distintas que resultam em menor capacidade de absorção de gordura pelas rosquinhas.

Esses resultados reforçam que TG4 se comporta de maneira significativamente diferente dos outros tratamentos, enquanto TG2, TG3 e TG1 podem ser considerados alternativas equivalentes em termos de absorção de gordura.

Isso é particularmente útil para estudos práticos, onde o foco é identificar tratamentos com comportamentos semelhantes ou distinguir aqueles que são inferiores.

Por fim, os resultados do teste de Scott-Knott fornecem uma visão objetiva e prática para a interpretação das diferenças entre os tratamentos. Com base nisso, pode-se sugerir que **TG4** seja investigado mais a fundo para entender as razões de sua menor performance, enquanto **TG2**, **TG3** e **TG1** podem ser usados de forma intercambiável, dependendo de fatores como custo ou disponibilidade.