

Analise TCC Milena

Carregar primeiro: tidyverse lawstat car readxl (linux package cmake)

```
library(tidyverse)
```

```
## — Attaching packages — tidyverse 1.3.2 —
## ✓ ggplot2 3.4.0      ✓ purrr 0.3.5
## ✓ tibble 3.1.8      ✓ dplyr 1.0.10
## ✓ tidyr 1.2.1       ✓ stringr 1.4.1
## ✓ readr 2.1.3       ✓ forcats 0.5.2
## — Conflicts — tidyverse_conflicts() —
## ✗ dplyr::filter() masks stats::filter()
## ✗ dplyr::lag() masks stats::lag()
```

```
library(lawstat)
library(car)
```

```
## Loading required package: carData
##
## Attaching package: 'car'
##
## The following object is masked from 'package:lawstat':
##
##     levene.test
##
## The following object is masked from 'package:dplyr':
##
##     recode
##
## The following object is masked from 'package:purrr':
##
##     some
```

Importar a planilha de dados

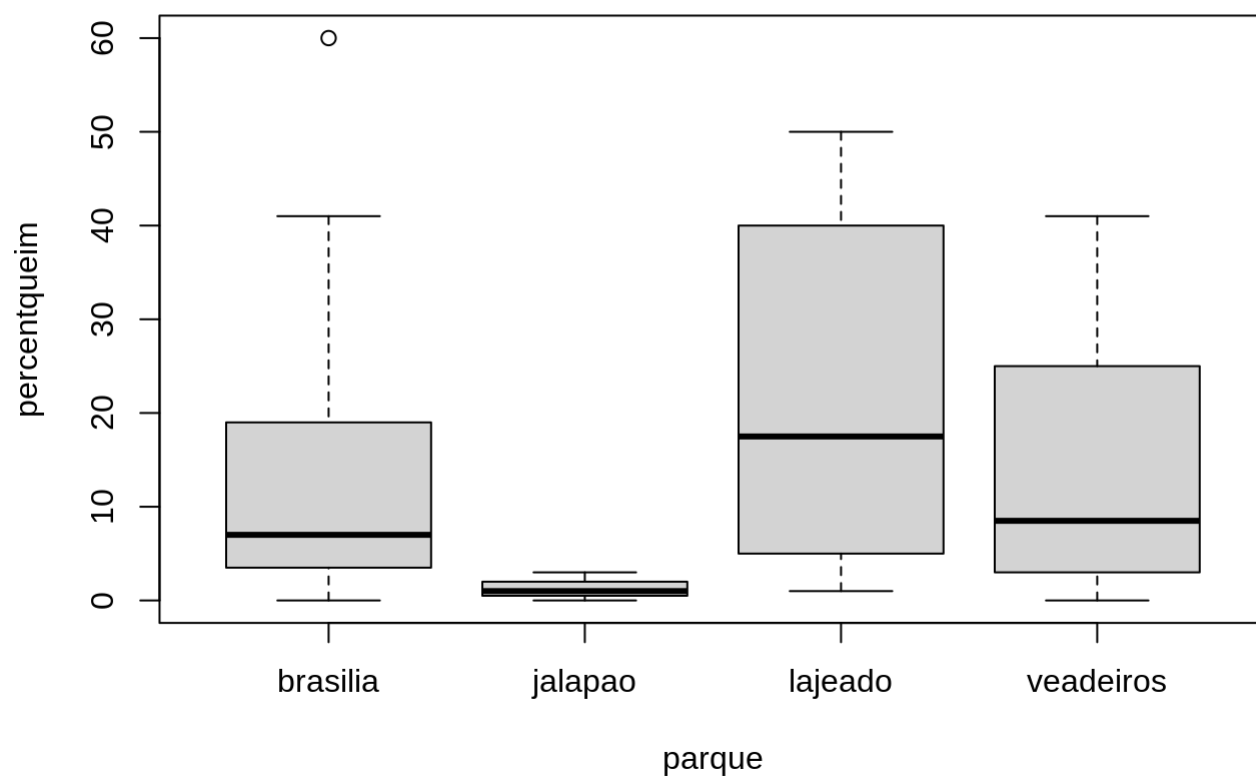
```
library(readxl)
nome_int_no_missing_pncv_corrigido <- read_excel("/home/rbcer10/rstudio-projects/pesq-
tcc-milena/dados/nome-int-no-missing-pncv-corrigido.xlsx")
```

Copiar a planilha importada para o dataframe parqueim

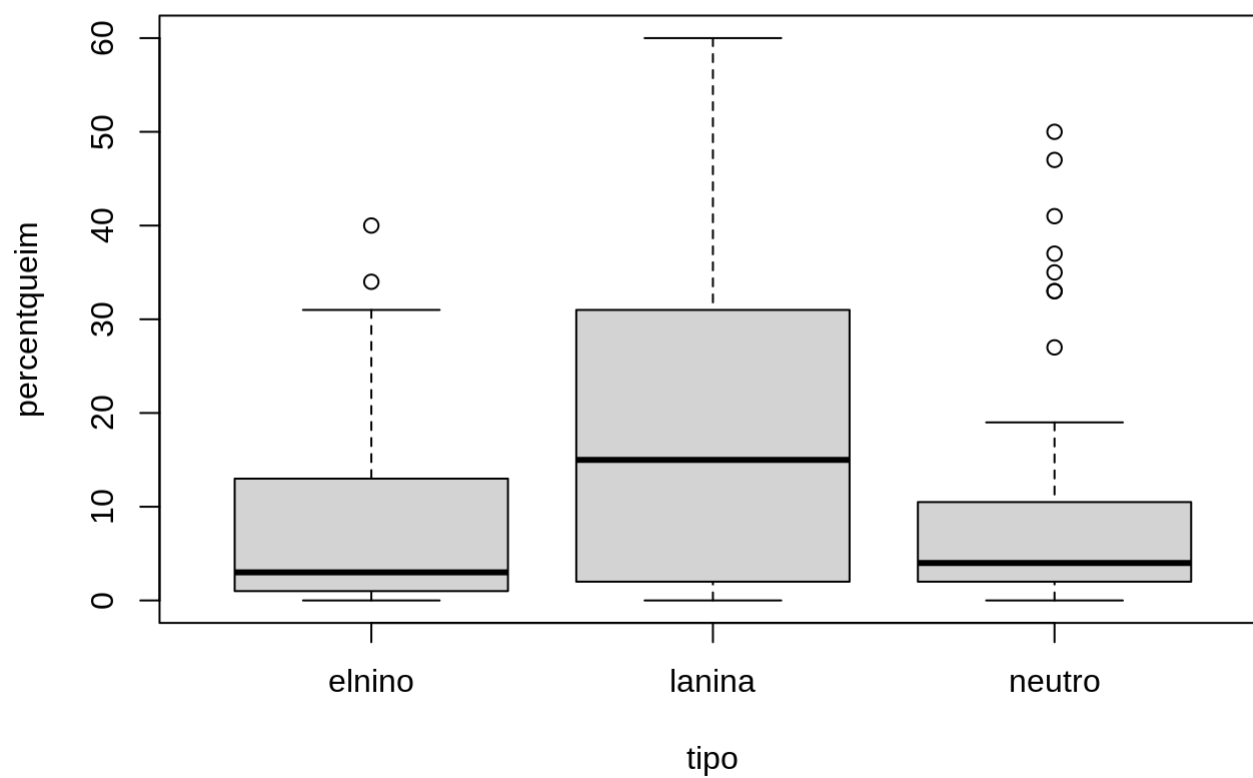
```
parqueim <- nome_int_no_missing_pncv_corrigido
```

Fazendo um boxplot dos dados olhando parques, int vs ext, e el nino

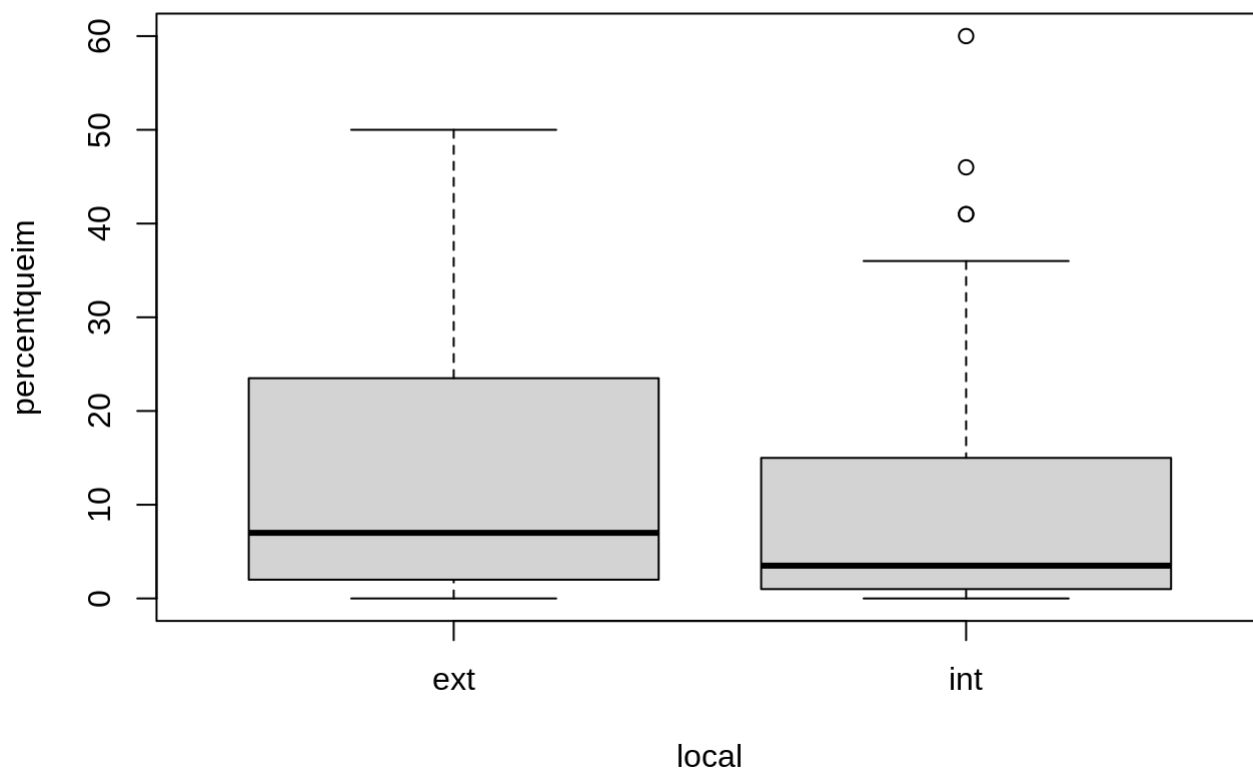
```
boxplot(percentqueim~parque, data=parqueim)
```



```
boxplot(percentqueim~tipo, data=parqueim)
```



```
boxplot(percentqueim~local, data=parqueim)
```



Interpretacao dos boxplots:

Percentagem queimada vs parques. O parque do jalapao queimou muito pouco e deve ser excluido da analise pois contribui pouco para a mesma. Os parques de Brasilia e Veadeiros tiveram padrao bem semelhante. O parque do Lajeado apresentou mediana mais alta e muita variancia, o que deve ter levado ao fato da anova a seguir ser quase significativa na variavel independente parques

Percentagem queimada vs clima. Os anos el nino e neutros tiveram medianas e variancias proximas entre si. Os anos la nina tiveram mediana e variancia bem mais alta, justificando o efeito significativo da variavel independente tipo na anova a seguir.

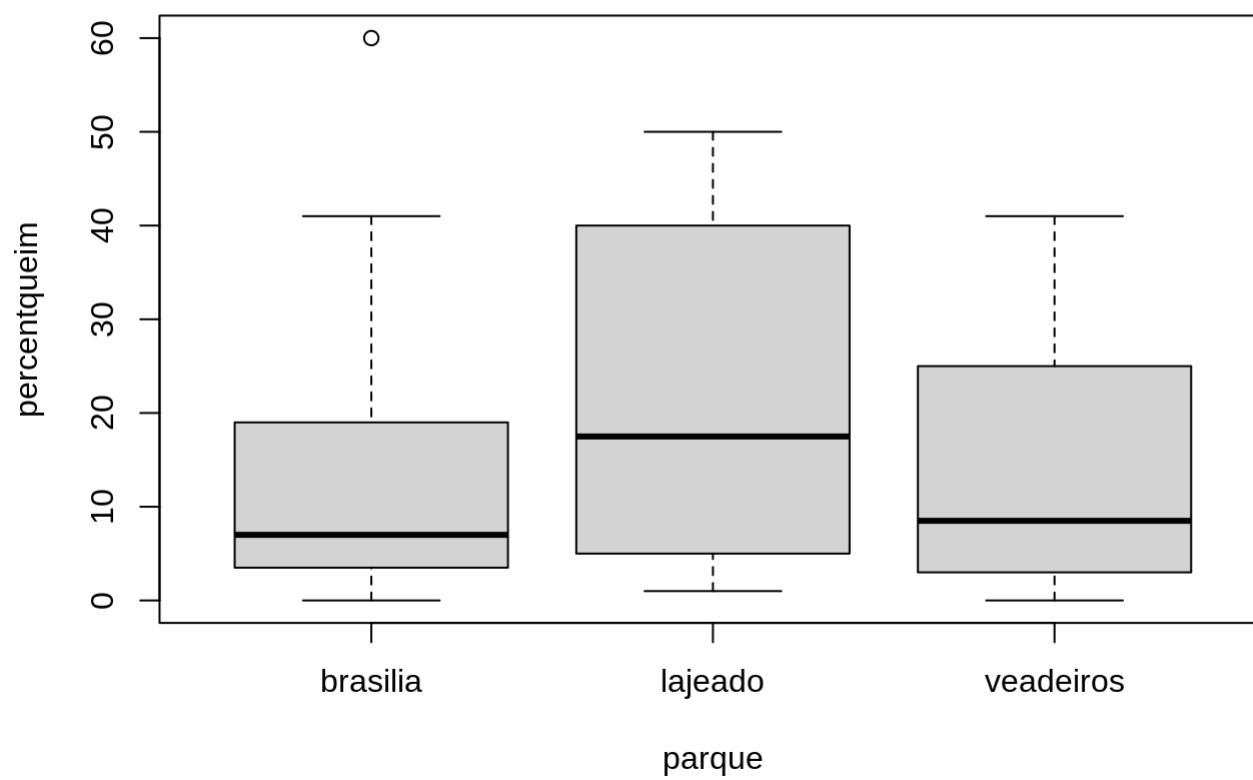
Percentagem queimada vs interno e externo. As areas externas tiveram mediana proxima das areas internas, mas as areas externas tiveram variancia muito maior, abrangendo toda a faixa de variacao das areas internas. A anova parques e local apresentou resultados nao significativos, e isto justificou a combinacao das areas externas e internas na analise global dos efeitos el nino e la nina

Removendo o parque do Jalapao da analise

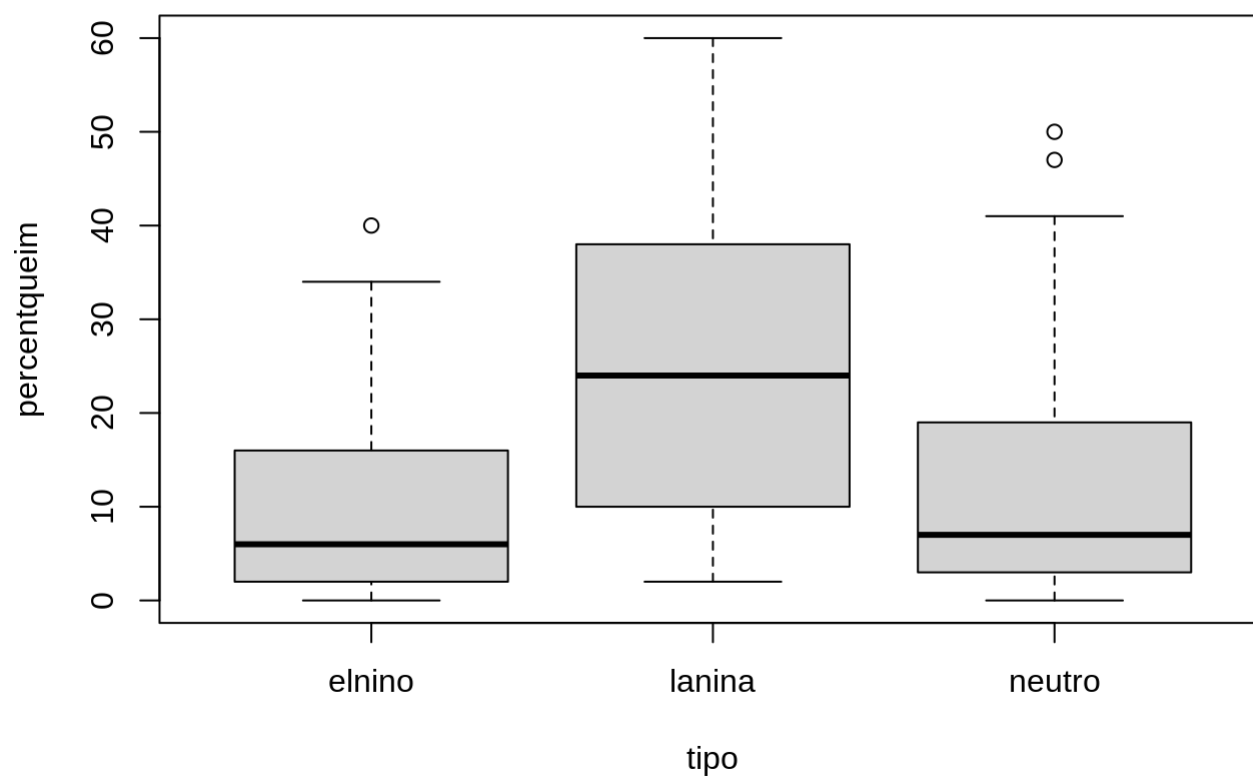
```
parquenj <- parqueim %>% filter(parque!='jalapao')
```

Repetindo o boxplot sem o jalapao:

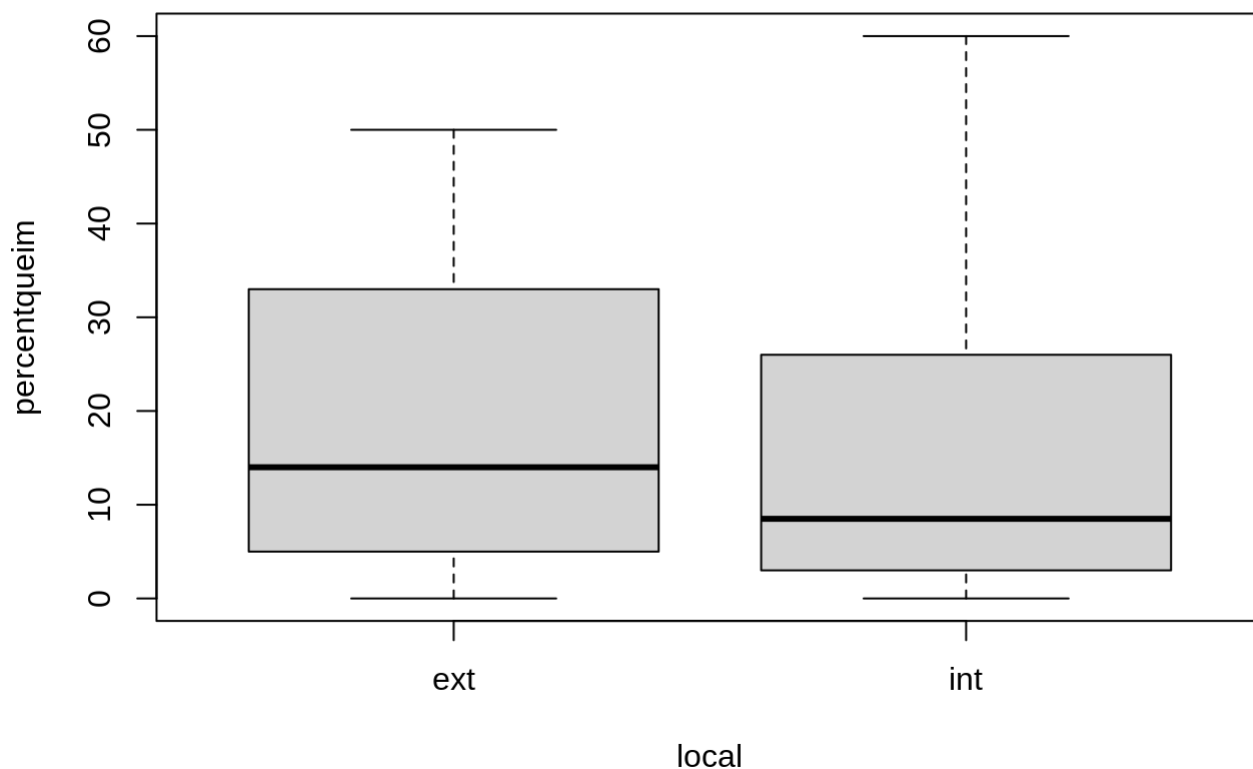
```
boxplot(percentqueim~parque, data=parquenj)
```



```
boxplot(percentqueim~tipo, data=parquenj)
```



```
boxplot(percentqueim~local, data=parquenj)
```



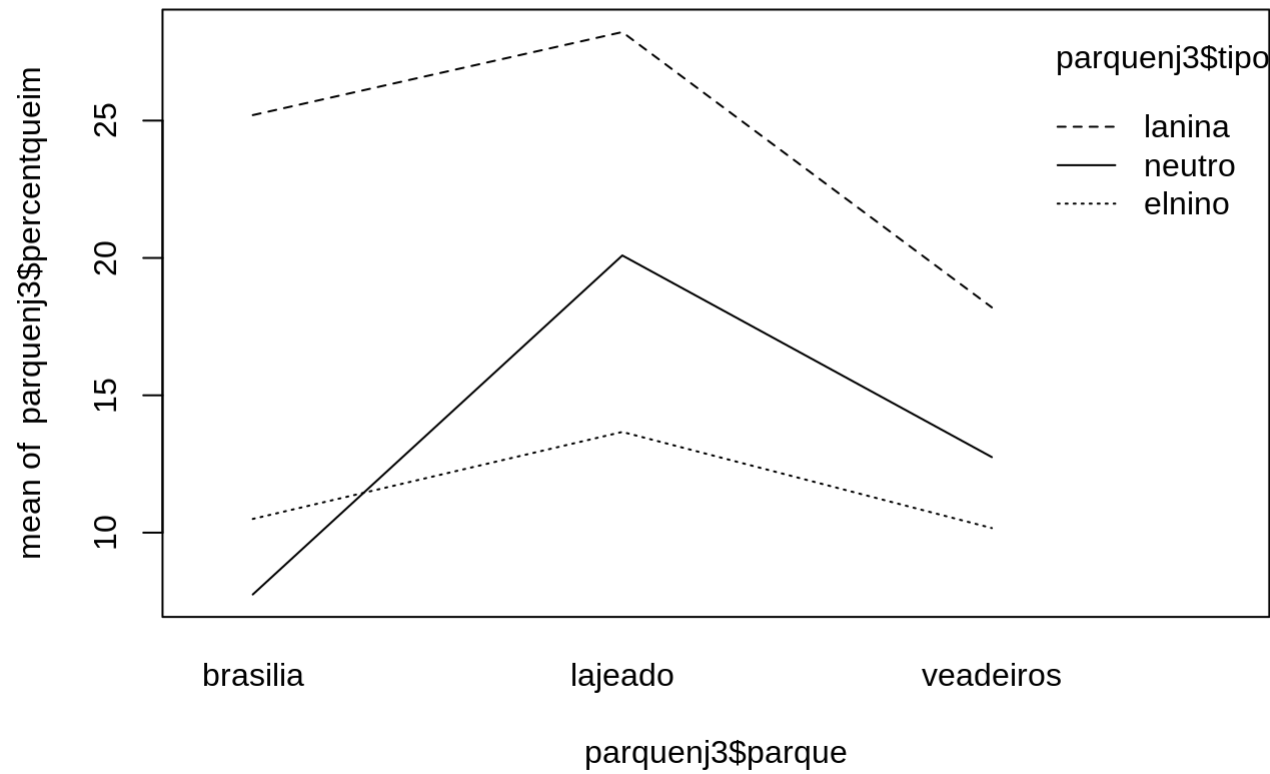
Fazendo transformacao logit da proporcao de area queimada cada parque cada ano e fazendo anova parque vs el nino com interacao e var dependente logit

O objetivo da transformacao logit e de normalizar os dados de proporcao para permitir o uso desta variavel em anova.

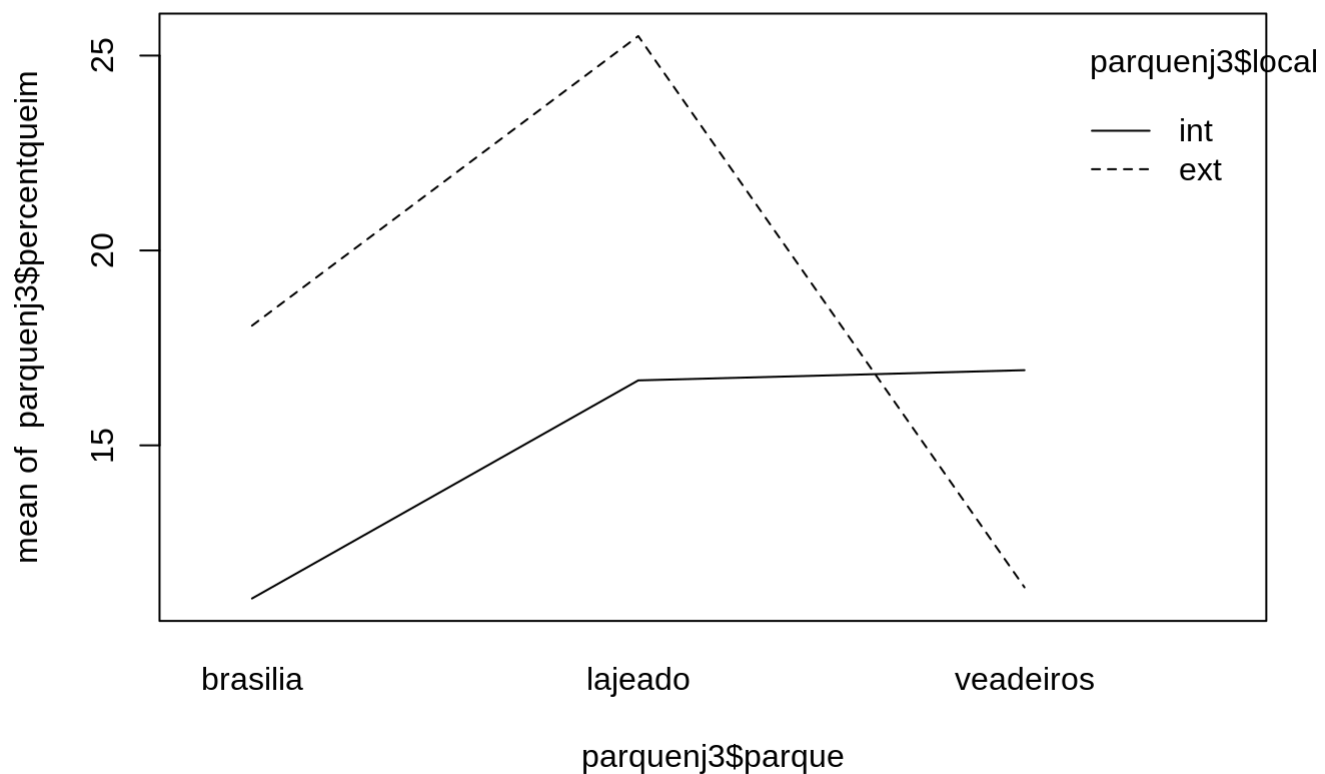
```
parquenj3 <- parquenj %>%
  mutate(logitquei = logit(propqueim))
anova3wtudo<-aov(logitquei ~ parque+tipo+parque:tipo, parquenj3)
summary(anova3wtudo)
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## parque      2   7.83    3.915    2.047 0.13641
## tipo        2  24.81   12.403    6.487 0.00255 **
## parque:tipo  4   0.49    0.123    0.065 0.99217
## Residuals   73 139.58    1.912
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
interaction.plot(parquenj3$parque, parquenj3$tipo, parquenj3$percentqueim)
```



```
interaction.plot(parquenj3$parque, parquenj3$local, parquenj3$percentqueim)
```

Interpretacao dos resultados:

ANOVA parque vs tipo

Media entre 3 parques na proporcao de area queimada tem diferenca nao significativa ($p < 0,09$, embora o efeito seja proximo de significativo. Media na proporcao de area queimada entre anos el nino, la nina, e neutro tem diferenca altamente significativa ($p < 0,007$). Conclusao: variancia entre parques nao significativa, variancia entre anos altamente significativa

Testando e confirmando a premissa de normalidade (shapiro) e homoscedascidade (levene) dos residuos. Os resultados sao nao significativos, portanto consistentes com a hipotese nula de que os dados nao diferem de de distribuicao normal e de distribuicoes com variancia homogenea

Grafico confirma este padrao mostrando que os anos la nina tem proporcao de area queimada muito mais alta que os anos el nino ou neutro

```
shapiro.test(resid(anova3wtudo))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  resid(anova3wtudo)
## W = 0.98077, p-value = 0.2565
```

```
leveneTest(parquenj3$logitquei, parquenj3$parque)
```

```
## Warning in leveneTest.default(parquenj3$logitquei, parquenj3$parque):
## parquenj3$parque coerced to factor.
```

	Df <int>	F value <dbl>	Pr(>F) <dbl>
group	2	0.3600592	0.6987739
	79	NA	NA
2 rows			

```
leveneTest(parquenj3$logitquei, parquenj3$tipo)
```

```
## Warning in leveneTest.default(parquenj3$logitquei, parquenj3$tipo):
## parquenj3$tipo coerced to factor.
```

	Df <int>	F value <dbl>	Pr(>F) <dbl>
group	2	1.688685	0.1913627
	79	NA	NA
2 rows			

Verificar se ha efeitos significativos interno vs externo

```
anova2wlocal<-aov(logitquei ~ parque+local+parque:local, parquenj3)
summary(anova2wlocal)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## parque      2   7.83   3.915   1.964 0.1474
## local       1   3.08   3.084   1.547 0.2174
## parque:local 2  10.28   5.141   2.579 0.0825 .
## Residuals  76 151.52   1.994
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Testando premissas de normalidade shapiro e homogeneidade de variancia levene dos residuos da ANOVA parques vs local Os resultados sao nao significativos, portanto consistentes com a hipotese nula de que os dados nao diferem de de distribuicao normal e de distribuicoes com variancia homogenea

```
shapiro.test(resid(anova2wlocal))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  resid(anova2wlocal)
## W = 0.97565, p-value = 0.1206
```

```
leveneTest(parquenj3$logitquei, parquenj3$local)
```

```
## Warning in leveneTest.default(parquenj3$logitquei, parquenj3$local):
## parquenj3$local coerced to factor.
```

	Df <int>	F value <dbl>	Pr(>F) <dbl>
group	1	0.413544	0.5220152
	80	NA	NA
2 rows			

CONCLUSAO GERAL

Considerando a variavel dependente proporcao de area queimada nos parques e seu entorno em cada ano, concluimos que o unico efeito significativo e referente ao clima, ou seja ha uma proporcao de area queimada significativamente mais alta nos anos la nina.

Ha algumas tendencias nao significativas em outras variaveis. As areas externas tem uma tendencia a queimar em maior proporcao e a variar mais entre anos do que as areas internas. Ha uma tendencia do parque do Lajeado de queimar em maior proporcao e de de variar mais entre anos do que os parques de Brasilia e da Chapada dos Veadeiros

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Ctrl+Shift+K* to preview the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.