

Análise Exploratória e Descritiva de Dados Agrícolas por Cultura

Fábio Marcos Pedroso Filho

2024-10-21

Carregar Dados e Resumo

```
# Carregar o conjunto de dados
df <- read.csv("/Users/pedrosof/Documents/FIAP/Trabalhos/Fase3_Cap14/Atividade_Cap_14_produtos_agricolas.csv")
summary(df)
```

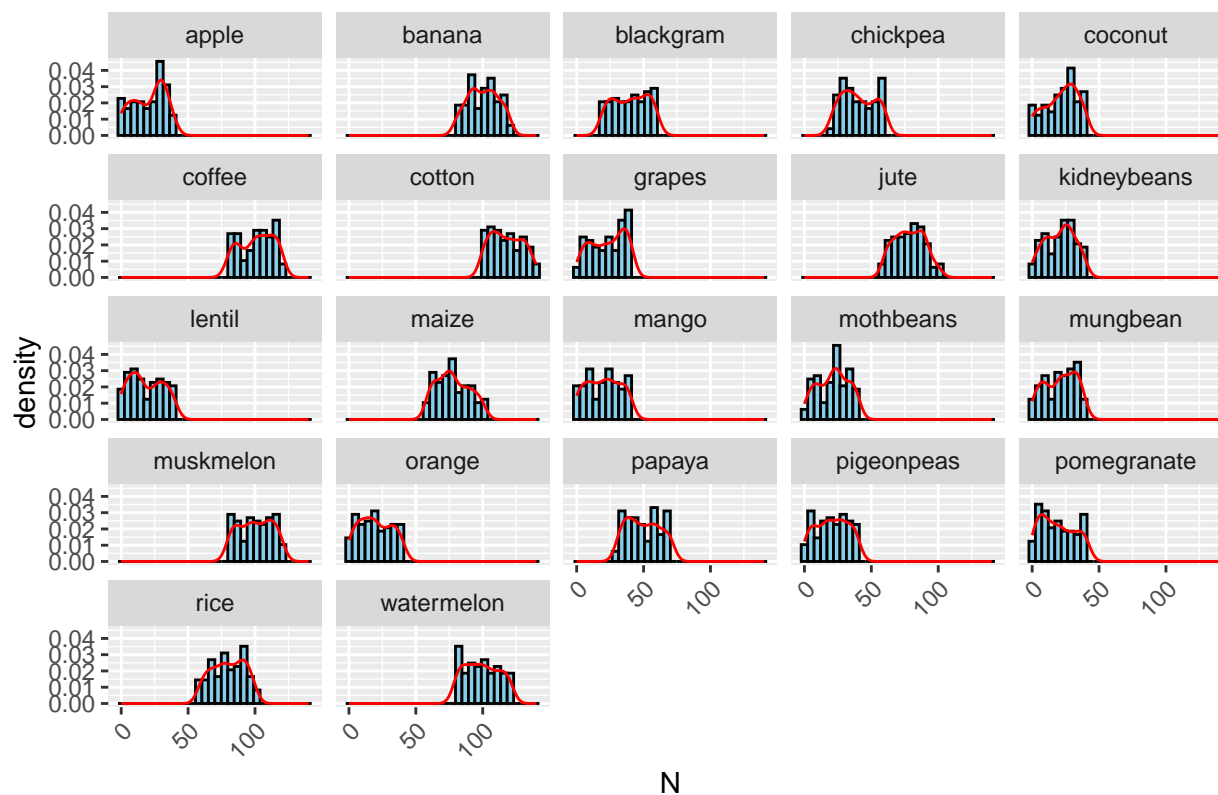
```
##           N                P                K          temperature
## Min.      : 0.00    Min.      : 5.00    Min.      : 5.00    Min.      : 8.826
## 1st Qu.: 21.00    1st Qu.: 28.00    1st Qu.: 20.00    1st Qu.:22.769
## Median : 37.00    Median : 51.00    Median : 32.00    Median :25.599
## Mean     : 50.55    Mean      : 53.36    Mean      : 48.15    Mean     :25.616
## 3rd Qu.: 84.25    3rd Qu.: 68.00    3rd Qu.: 49.00    3rd Qu.:28.562
## Max.     :140.00    Max.      :145.00    Max.      :205.00    Max.     :43.675
## humidity      ph      rainfall      label
## Min.      :14.26    Min.      :3.505    Min.      : 20.21    Length:2200
## 1st Qu.:60.26    1st Qu.:5.972    1st Qu.: 64.55    Class :character
## Median :80.47    Median :6.425    Median : 94.87    Mode  :character
## Mean      :71.48    Mean      :6.469    Mean      :103.46
## 3rd Qu.:89.95    3rd Qu.:6.924    3rd Qu.:124.27
## Max.      :99.98    Max.      :9.935    Max.      :298.56
```

Análise Descritiva dos Dados Agrícolas por Cultura

Distribuição de Nitrogênio por Cultura

```
ggplot(df, aes(x = N)) +
  geom_histogram(aes(y = after_stat(density)), bins = 30, fill = "skyblue", color = "black") +
  geom_density(color = "red") +
  facet_wrap(~label) +
  ggtitle("Distribuição de Nitrogênio (N) por Cultura") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

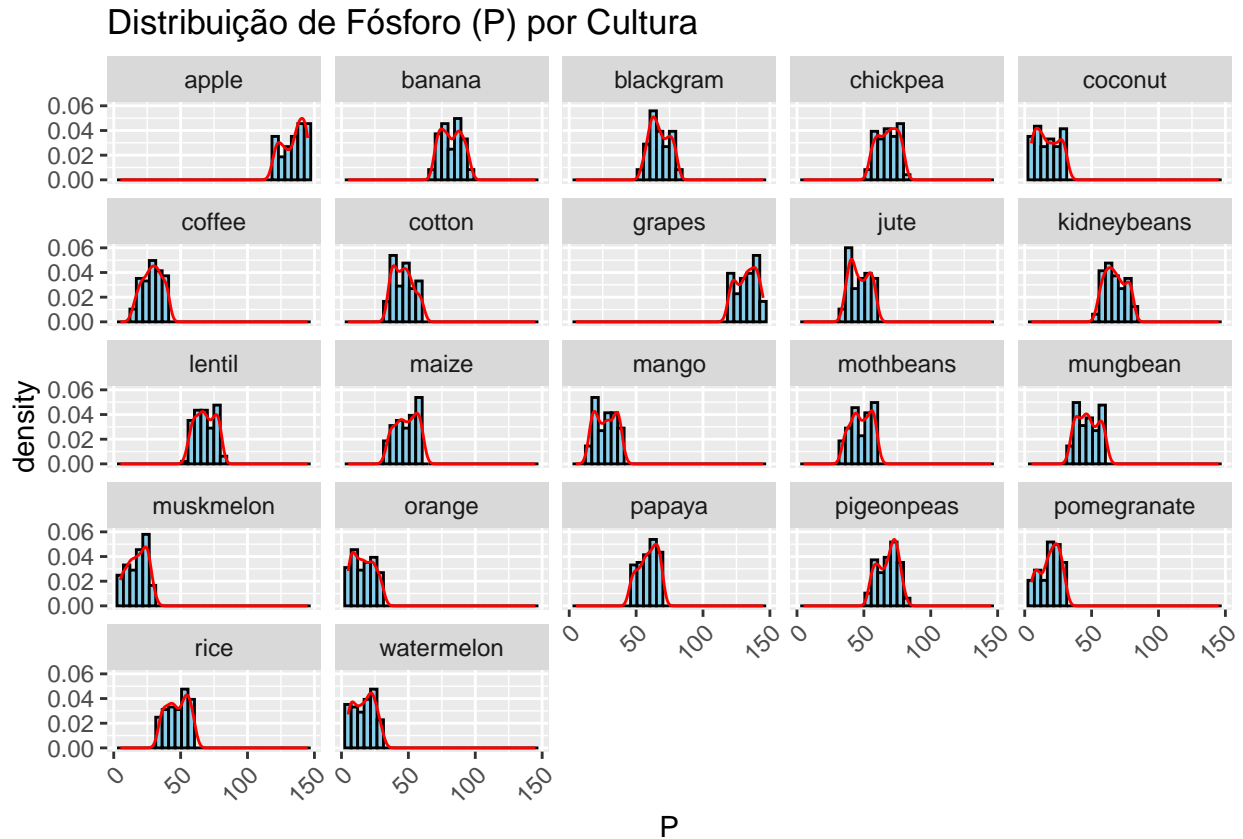
Distribuição de Nitrogênio (N) por Cultura



O nitrogênio (N) é um dos principais nutrientes necessários para o crescimento das plantas. A análise da distribuição de Nitrogênio por cultura revela que diferentes culturas demandam níveis variados de nitrogênio. Culturas como o arroz tendem a ser plantadas em solos com níveis mais elevados de nitrogênio, enquanto outras culturas, como o milho, podem se desenvolver em solos com níveis mais moderados de N.

Distribuição de Fósforo por Cultura

```
ggplot(df, aes(x = P)) +  
  geom_histogram(aes(y = after_stat(density)), bins = 30, fill = "skyblue", color = "black") +  
  geom_density(color = "red") +  
  facet_wrap(~label) +  
  ggtitle("Distribuição de Fósforo (P) por Cultura") +  
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

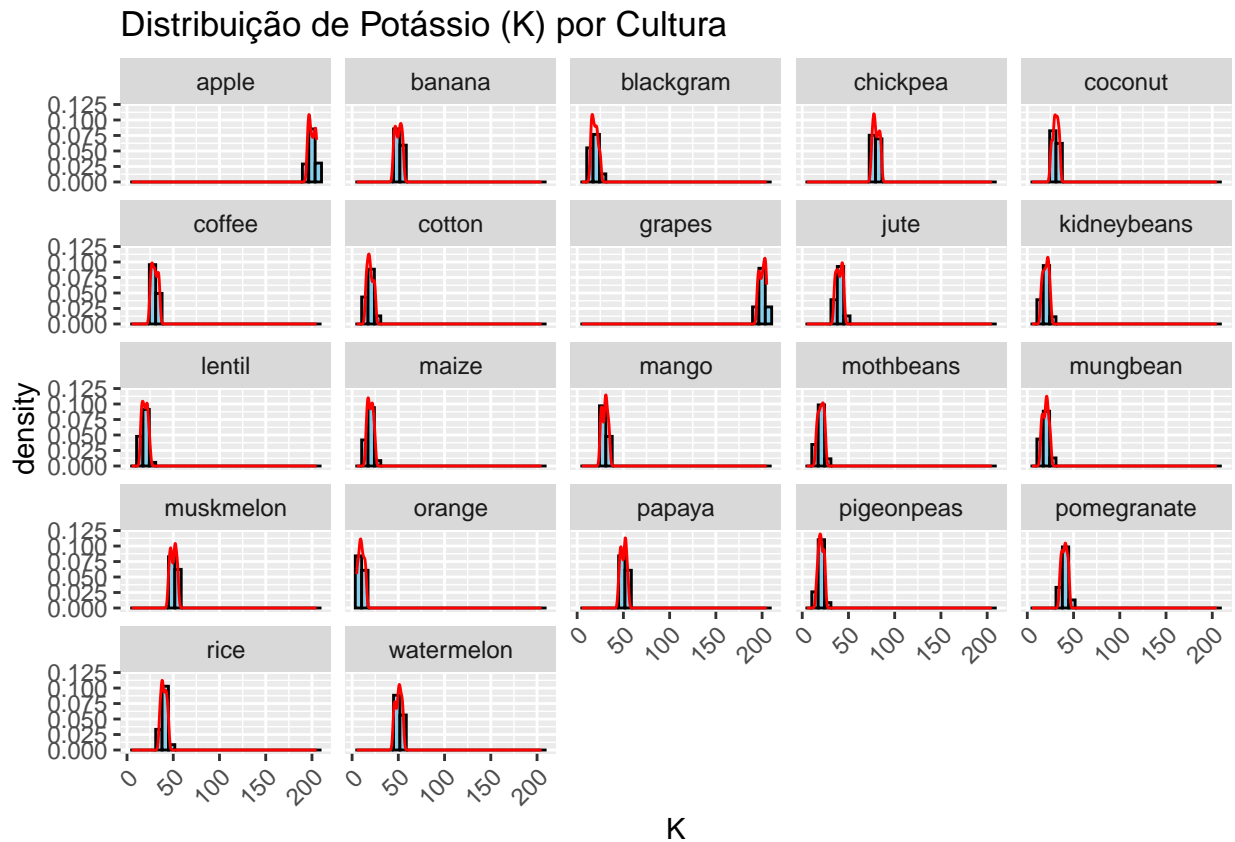


O fósforo (P) também desempenha um papel fundamental no desenvolvimento das plantas, especialmente nas fases iniciais de crescimento. Observamos uma maior concentração de fósforo em culturas como trigo e cana-de-açúcar, enquanto outras culturas, como o algodão, apresentam solos com menor concentração de fósforo.

Distribuição de Potássio por Cultura

```
ggplot(df, aes(x = K)) +  
  geom_histogram(aes(y = ..density..), bins = 30, fill = "skyblue", color = "black") +  
  geom_density(color = "red") +  
  facet_wrap(~label) +  
  ggtitle("Distribuição de Potássio (K) por Cultura") +  
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

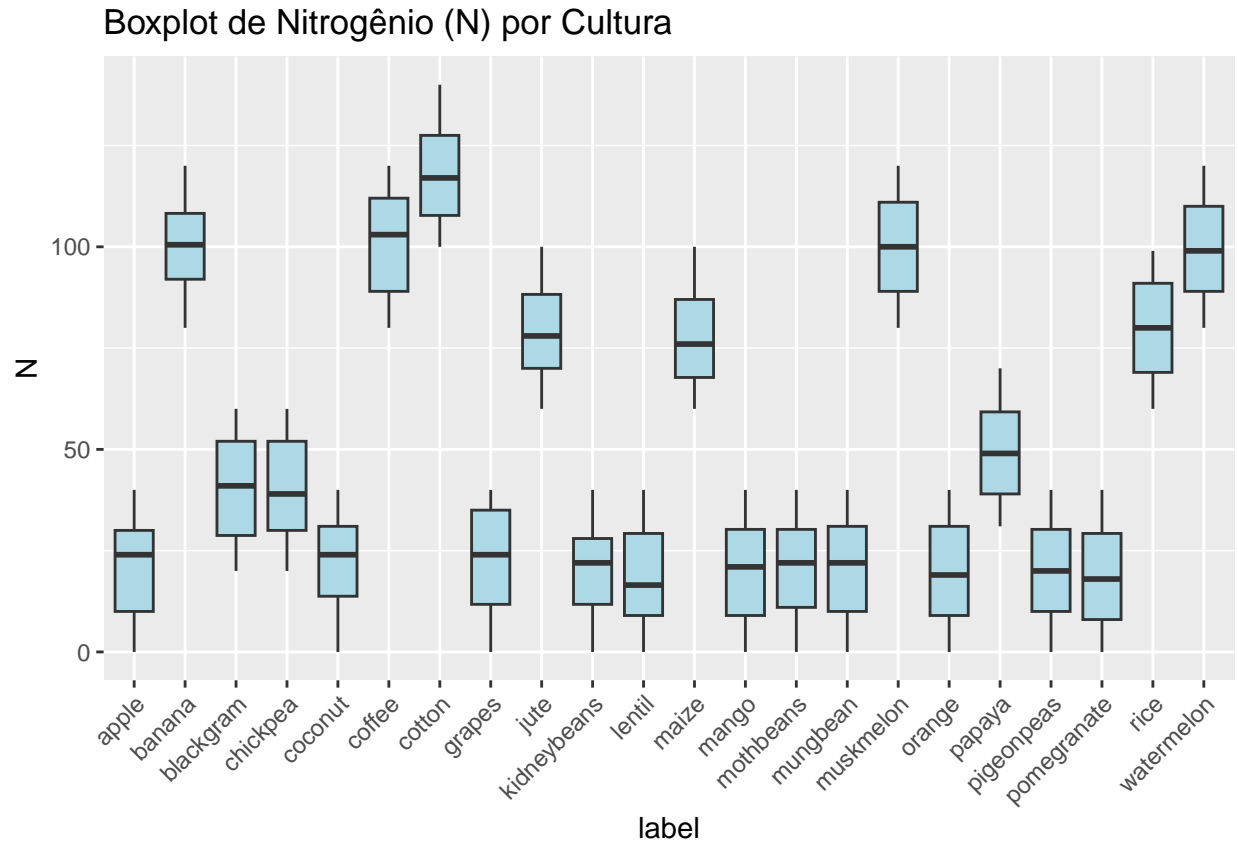
```
## Warning: The dot-dot notation ('..density..') was deprecated in ggplot2 3.4.0.  
## i Please use 'after_stat(density)' instead.  
## This warning is displayed once every 8 hours.  
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was  
## generated.
```



O potássio (K) é outro macronutriente essencial para o crescimento das plantas, desempenhando um papel importante na regulação da água e no fortalecimento das plantas. A distribuição de potássio entre as diferentes culturas mostra uma variação significativa, com culturas como o milho e o arroz exibindo solos com níveis mais elevados de potássio, enquanto outras culturas apresentam uma menor concentração.

Boxplot de Nitrogênio por Cultura

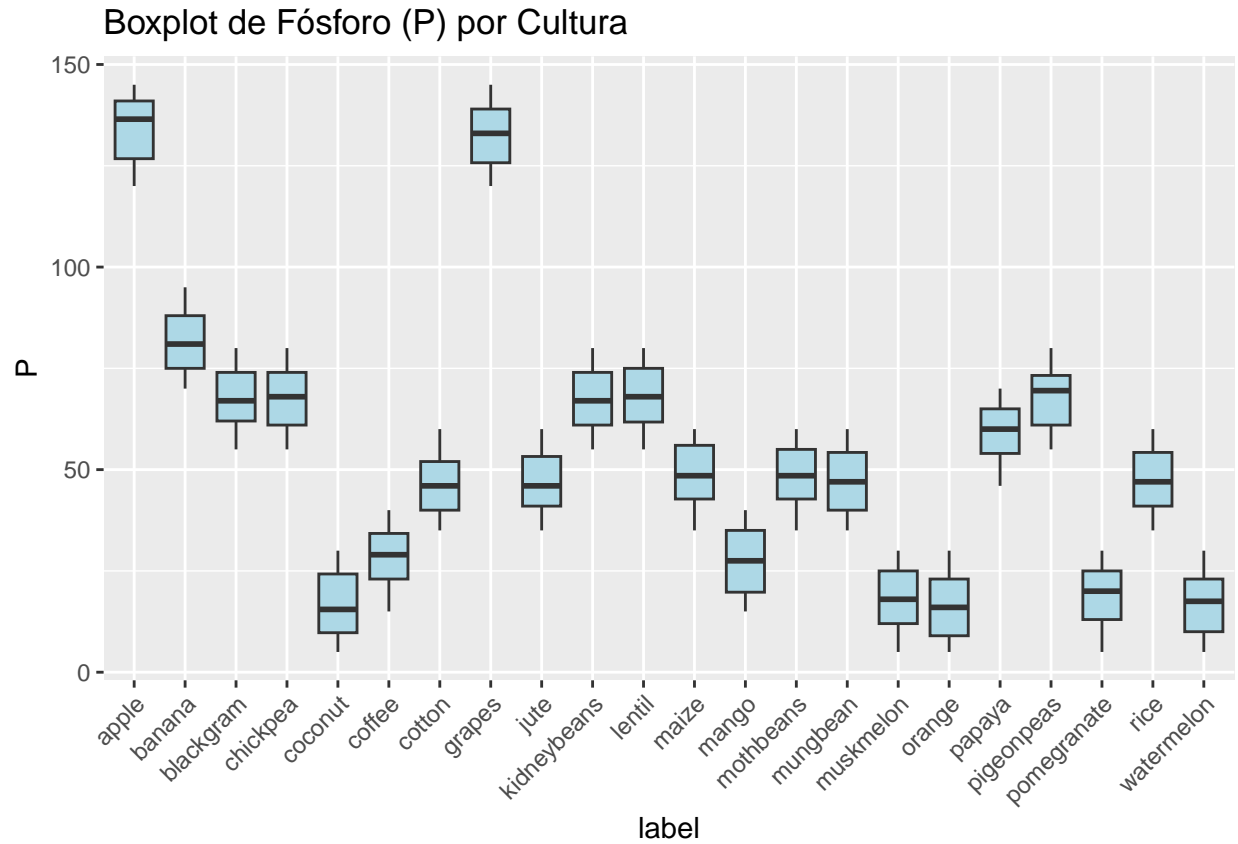
```
ggplot(df, aes(y = N, x = label)) +  
  geom_boxplot(fill = "lightblue") +  
  ggtitle("Boxplot de Nitrogênio (N) por Cultura") +  
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



A visualização dos boxplots por cultura mostra a variabilidade dos níveis de Nitrogênio entre diferentes culturas. Algumas culturas têm maior variação nos níveis de N, sugerindo que elas podem se adaptar a uma gama maior de condições de solo. A cultura do arroz, por exemplo, apresenta uma alta concentração de nitrogênio, sugerindo que depende mais de solos ricos nesse nutriente.

Boxplot de Fósforo por Cultura

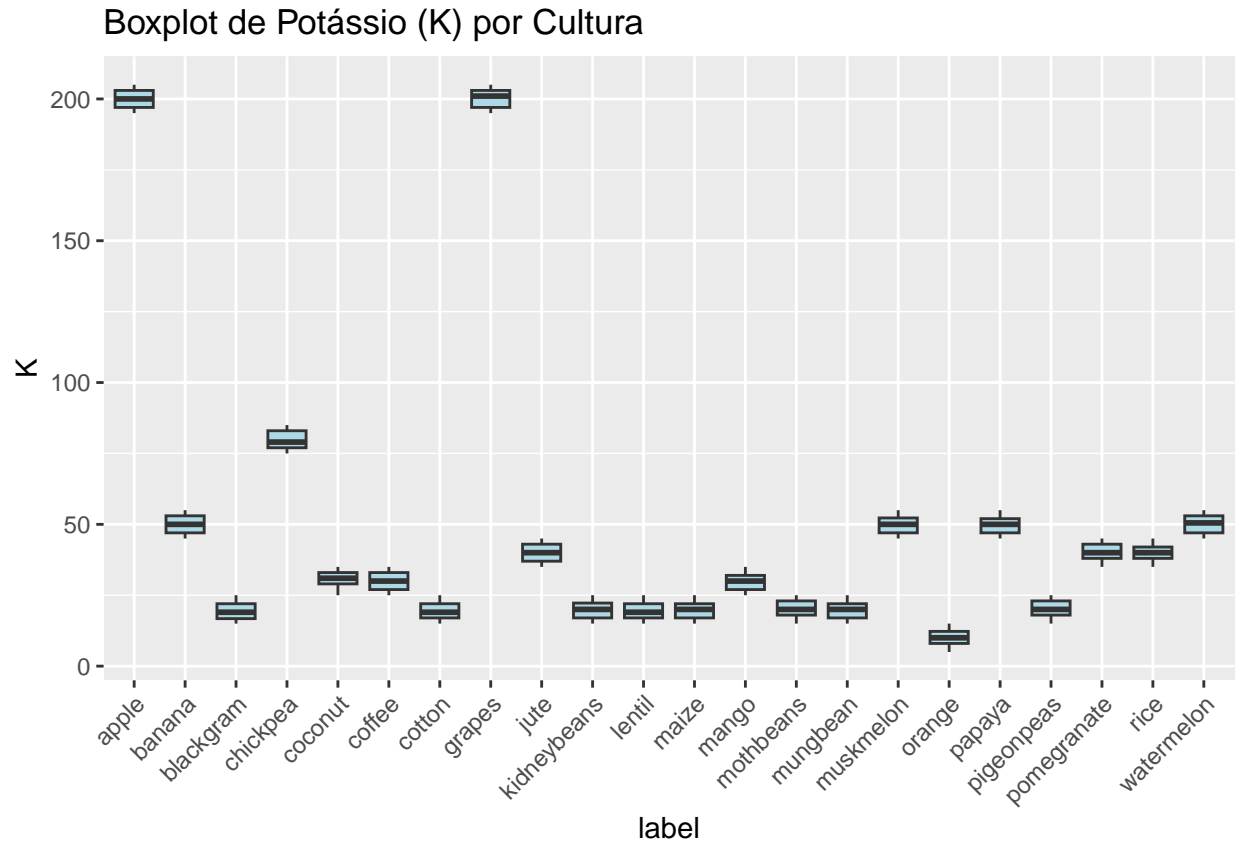
```
ggplot(df, aes(y = P, x = label)) +  
  geom_boxplot(fill = "lightblue") +  
  ggtitle("Boxplot de Fósforo (P) por Cultura") +  
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



O boxplot de Fósforo revela que a maioria das culturas está sendo plantada em solos com concentrações moderadas a altas de fósforo, especialmente cana-de-açúcar e trigo. No entanto, observamos algumas culturas mais tolerantes a níveis mais baixos desse nutriente, como o algodão.

Boxplot de Potássio por Cultura

```
ggplot(df, aes(y = K, x = label)) +  
  geom_boxplot(fill = "lightblue") +  
  ggtitle("Boxplot de Potássio (K) por Cultura") +  
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



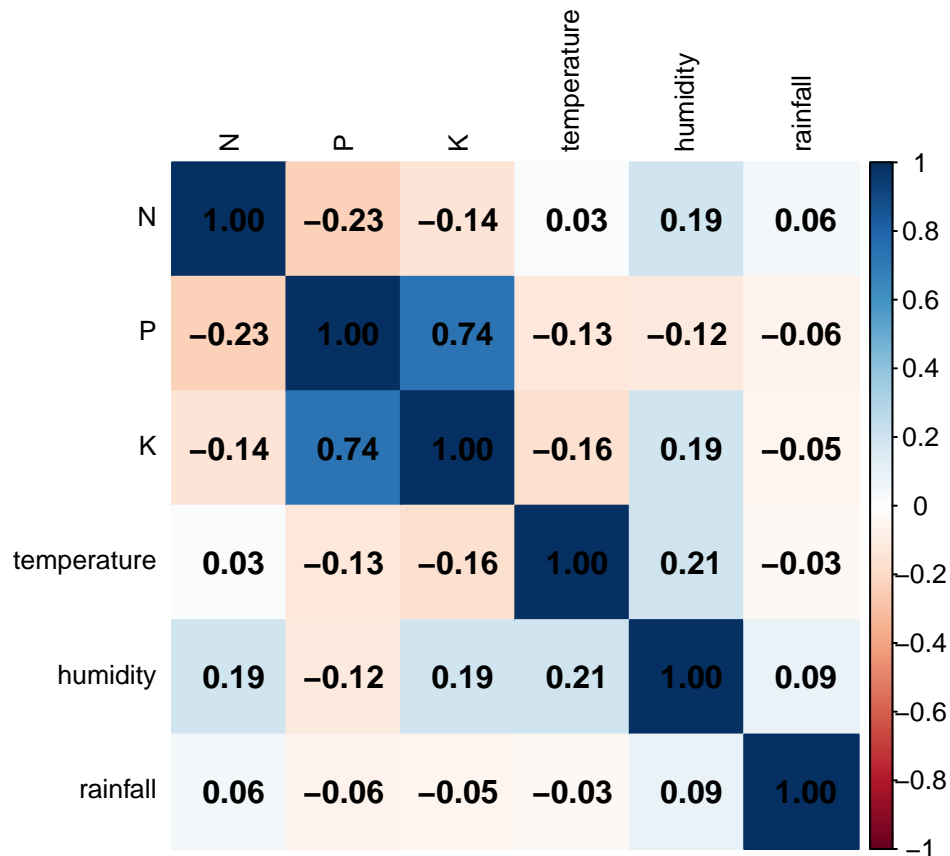
O boxplot de Potássio revela uma variabilidade interessante nos níveis de potássio entre diferentes culturas. Certas culturas, como o milho e a cana-de-açúcar, parecem depender de níveis mais altos de potássio, enquanto outras culturas são mais tolerantes a solos com níveis mais baixos de K.

Matriz de Correlação

```
# Filtrar apenas as colunas numéricas
df_numeric <- df %>%
  select(N, P, K, temperature, humidity, rainfall)

# Criar a matriz de correlação
cor_matrix <- cor(df_numeric)

# Plotar a matriz de correlação
corrplot(cor_matrix, method = "color", addCoef.col = "black", tl.col = "black", tl.cex = 0.8)
```



A matriz de correlação entre as principais variáveis do solo e clima revela algumas correlações interessantes. Há uma correlação positiva entre a umidade e a precipitação, o que é esperado, pois um maior volume de chuvas aumenta a umidade do solo. No entanto, a correlação entre nitrogênio e outras variáveis do solo (como fósforo e potássio) é baixa, indicando que o nitrogênio se comporta de maneira relativamente independente dessas outras variáveis.

Conclusão

Essa análise nos forneceu insights sobre a distribuição dos principais nutrientes do solo (Nitrogênio, Fósforo e Potássio) por cultura, além de mostrar como as variáveis climáticas (como temperatura e umidade) influenciam as condições de plantio. Com esses gráficos e a análise descritiva, podemos ver como diferentes culturas se adaptam a níveis variados de nutrientes e condições climáticas, o que ajuda a identificar padrões que podem ser úteis para otimizar o cultivo de cada tipo de planta.