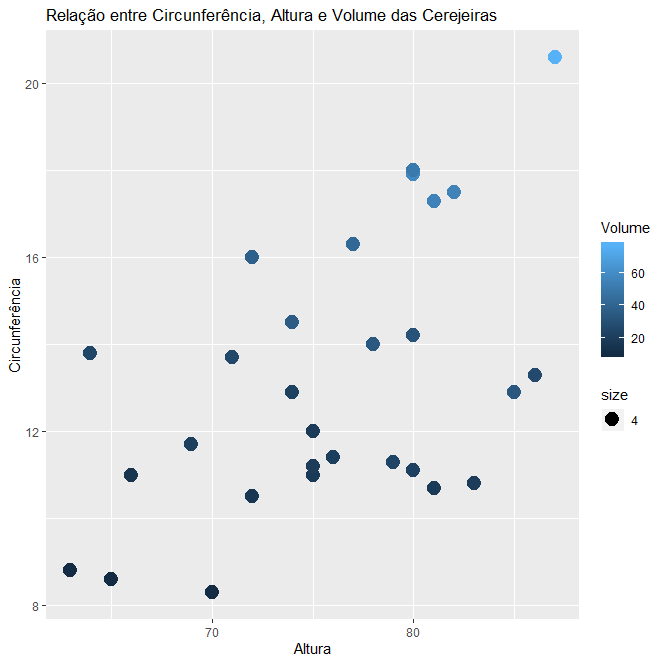
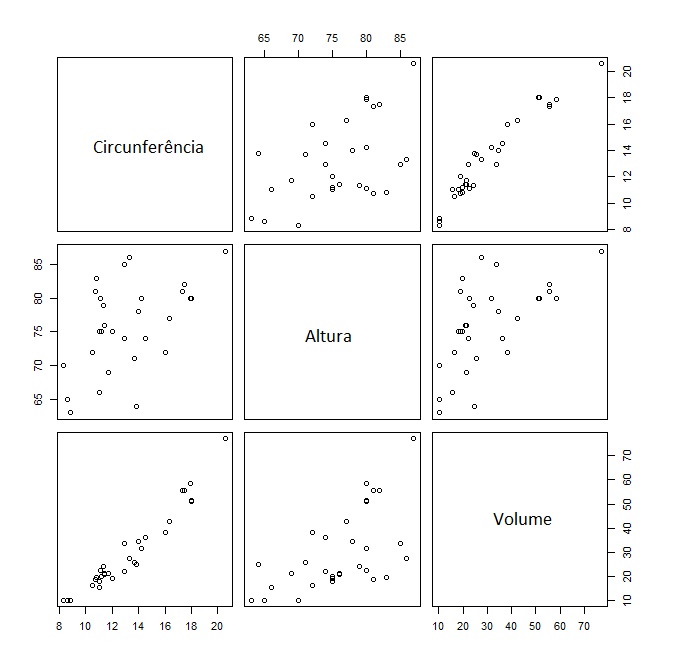
Atividade de Estatística no R

1. A atividade 1 foi produzir um gráfico no R utilizando os dados disponíveis no software chamado ‘trees’. Os dados correspondem a uma tabela contendo as informações sobre circunferência, altura e volume de 31 cerejeiras. O objetivo da análise estatística foi observar a relação entre a circunferência, a altura e o volume das cerejeiras. Desse modo, foi utilizada a função ggplot para determinar como essas três variáveis estavam relacionadas entre si.

Na imagem exportada do R após a análise estatística, pode ser observado que à medida que a circunferência das cerejeiras aumenta, o volume e a altura também aumentam. Desse modo podemos dizer que essas variáveis são diretamente proporcionais. A relação diretamente proporcional entre circunferência, altura e volume faz sentido. Quanto mais alta é a planta, maior deve ser sua circunferência para pode sustentar o peso do tronco, pois de outro modo, a planta poderia quebrar o caule com muita facilidade. Quanto maior a circunferência e a altura de um tronco de cone (elemento geométrico, fórmula matemática V = π.h/3 . (R2 + R . r +r2), maior o volume.



Na figura 2 é possível observar a como as variáveis estão correlacionadas entre si, par a par utilizando a função plot (trees).



2. A atividade 2 foi realizada utilizando os dados “Soils” disponíveis no R. Os dados de ‘soils’ correspondem a 52 amostras de solo retiradas de 4 blocos de solo, as quais foram coletadas em profundidades diferentes. As variáveis analisadas foram tipo de solo, ph, cálcio, potássio, condutividade, sódio, magnésio, nitrogênio e densidade. O objetivo da análise estatística foi construir um modelo linear com mais de dois preditores, determinar como o pH está relacionado as outras variáveis e se os preditores são ou não significativos para explicar os dados do modelo.

O sumário do modelo abaixo mostra com asterisco as variáveis que foram significativas no modelo.

> summary(m2\_glm)

Call:

glm(formula = pH ~ Block + N + Dens + Ca + Mg + K + Na + Conduc,

data = Soils)

Deviance Residuals:

Min 1Q Median 3Q Max

-0.68129 -0.13257 0.02164 0.13648 1.09103

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4.205545 0.902591 4.659 4.02e-05 \*\*\*

Block2 0.502166 0.151374 3.317 0.00205 \*\*

Block3 0.216036 0.160264 1.348 0.18585

Block4 0.495536 0.141845 3.494 0.00125 \*\*

N -2.154734 1.856766 -1.160 0.25329

Dens 0.061610 0.464283 0.133 0.89515

Ca 0.159556 0.035170 4.537 5.84e-05 \*\*\*

Mg -0.001083 0.048258 -0.022 0.98222

K -0.848877 0.396262 -2.142 0.03882 \*

Na 0.093890 0.078777 1.192 0.24091

Conduc -0.167404 0.066376 -2.522 0.01610 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for gaussian family taken to be 0.1014823)

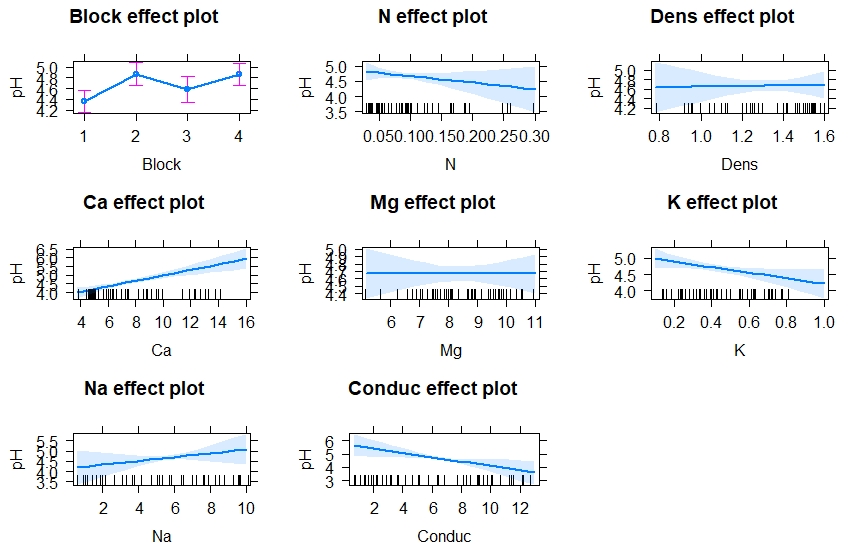
Null deviance: 21.2153 on 47 degrees of freedom

Residual deviance: 3.7548 on 37 degrees of freedom

AIC: 37.907

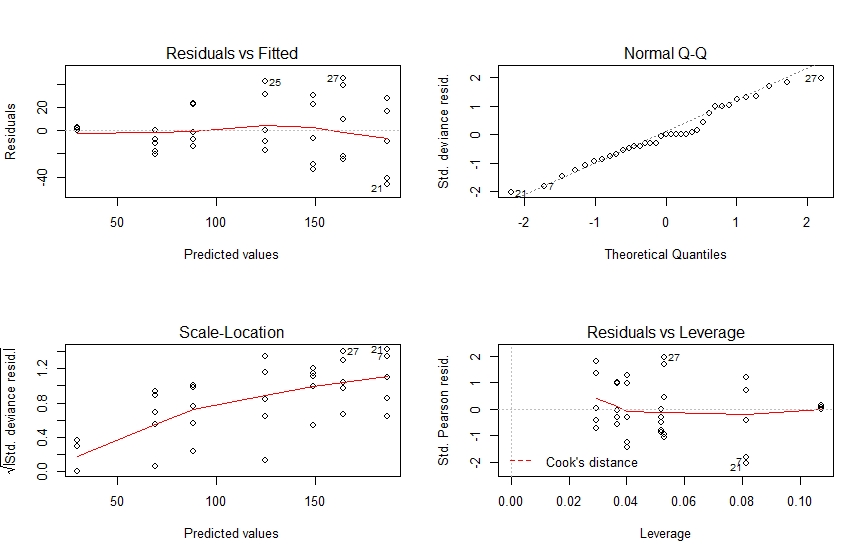
Number of Fisher Scoring iterations: 2

~~As variáveis bloco do solo não puderam ser analisadas nos modelos lineares por serem categóricas~~. Os preditores significativos para explicar o pH do solo foram Cálcio (Ca), Potássio (K) e Condutividade. O pH foi diretamente proporcional ao cálcio e inversamente proporcional ao potássio e à condutividade como é possível observar no plot dos efeitos do modelo, pacote ‘effects’.



O cálcio geralmente é aplicado no solo ácido para elevar o pH. (ABREU JUNIOR *et al*., 2000). Assim quanto mais cálcio, maior o pH do solo. A condutividade do solo aumenta quanto mais íons estão disponíveis. Num solo ácido, muitos íons hidrogênio estão no meio, o que aumenta a condutividade. A correlação entre pH e potássio no solo é conhecida pelos agrônomos, quanto mais ácido o solo, menor o teor de íons potássio (CATANI; KUPPER, 1946).

O modelo atendeu aos pressupostos e pode ser utilizado conforme indicado na análise dos gráficos abaixo, os quais foram obtidos usando a função plot do modelo (plot(m2\_glm).



1. Script da Atividade 1.

> trees

> library(ggplot2)

> g<-ggplot(aes(y=Girth, x=Height, color=Volume,size=4),data=trees)+ ggtitle("Relação entre Circunferência, Altura e Volume")

> g+geom\_point()+xlab("Altura")+ylab("Circunferência")+ ggtitle("Relação entre Circunferência, Altura e Volume das Cerejeiras")

> plot(trees)

B) Script Atividade 2

> help("Soils")

> Soils

> m2\_glm <- glm(Block ~ pH, data=Soils)

> m2\_glm <- glm(pH ~Block+N+Dens+Ca+Mg+K+Na+Conduc, data=Soils)

> summary(m2\_glm)

> par(mfrow=c(2,2))

> plot(m2\_glm)

>summary(glht(m2\_glm,linfect=mcp(Block+N+Dens+Ca+Mg+K+Na+Conduc="Tukey")))

> confint(m2\_glm)

> plot(allEffects(m2\_glm))

Referências Bibliográficas

ABREU JUNIOR, C. H. *et al* . Condutividade elétrica, reação do solo e acidez potencial em solos adubados com composto de lixo.**Rev. Bras. Ciênc. Solo**,  Viçosa ,  v. 24, n. 3, p. 635-647,  Sept.  2000 .   Disponível em <http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-06832000000300016&lng=en&nrm=iso>. Acesso em  07  Feb.  2020.  <http://dx.doi.org/10.1590/S0100-06832000000300016>.

CATANI, R.A.; KUPPER, A. Algumas características químicas dos solos do Estado de São Paulo e sua interpretação analítica. Bragantina, v. 6, p. 147-163, 1946. Disponível em <<http://www.scielo.br/pdf/brag/v6n4/01.pdf>> Acesso em 07 de fev. 2020.