

Curso Livre II

Aluno: Manuel Ferreira Junior

Matricula: 20180008601

Modulo I

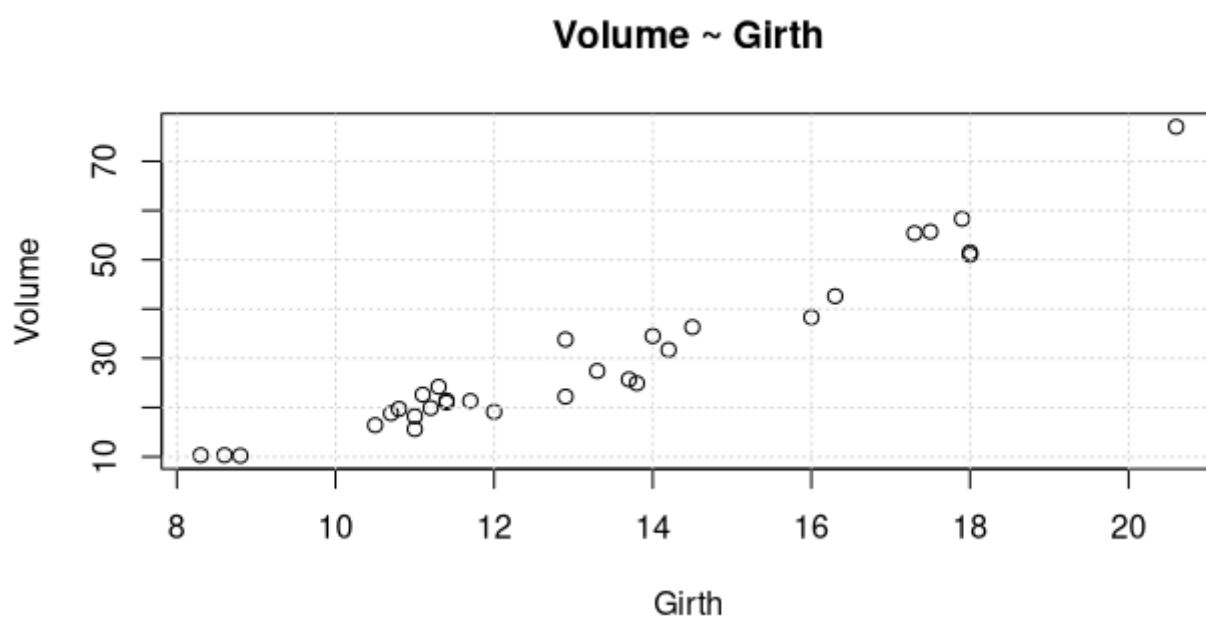
Tarefa 01 Acesse o banco de dados trees no R, utilizando a função data(). Estes dados fornecem as informações sobre o diâmetro (Girth), a altura (Height) e o volume (Volume) de 31 árvores cerejeiras. Faça uma análise sobre a relação linear entre as variáveis Girth e Volume e ajuste um modelo de regressão linear simples para explicar o volume da árvore em função do seu diâmetro. E depois responda as seguintes questões:

Carregando os dados

```
data(trees)
attach(trees)
```

A) Obtenha o gráfico de dispersão. Como você descreveria a relação entre as duas variáveis?

```
> plot(Girth,Volume)
> grid()
> title('Volume ~ Girth')
```



B) Qual o grau de correlação linear entre as variáveis?

```
> cor(Girth, Volume)
[1] 0.9671194
```

C) Qual a conclusão do teste para o coeficiente de correlação? Justifique utilizando o p-valor.

```
> cor.test(Volume, Girth)

Pearson's product-moment correlation

data:  Volume and Girth
t = 20.478, df = 29, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.9322519 0.9841887
sample estimates:
      cor
0.9671194
```

D) Ajuste um modelo de regressão linear simples.

```
> gvmmodel <- lm(Volume ~ Girth)
> gvmmodel

Call:
lm(formula = Volume ~ Girth)

Coefficients:
(Intercept)      Girth
   -36.943       5.066
```

E) Qual a variável resposta (interesse) e a regressora (explicativa)?

Para esse modelo, consideremos que a variável explicativa é o diâmetro da árvore e a variável resposta é o volume da árvore.

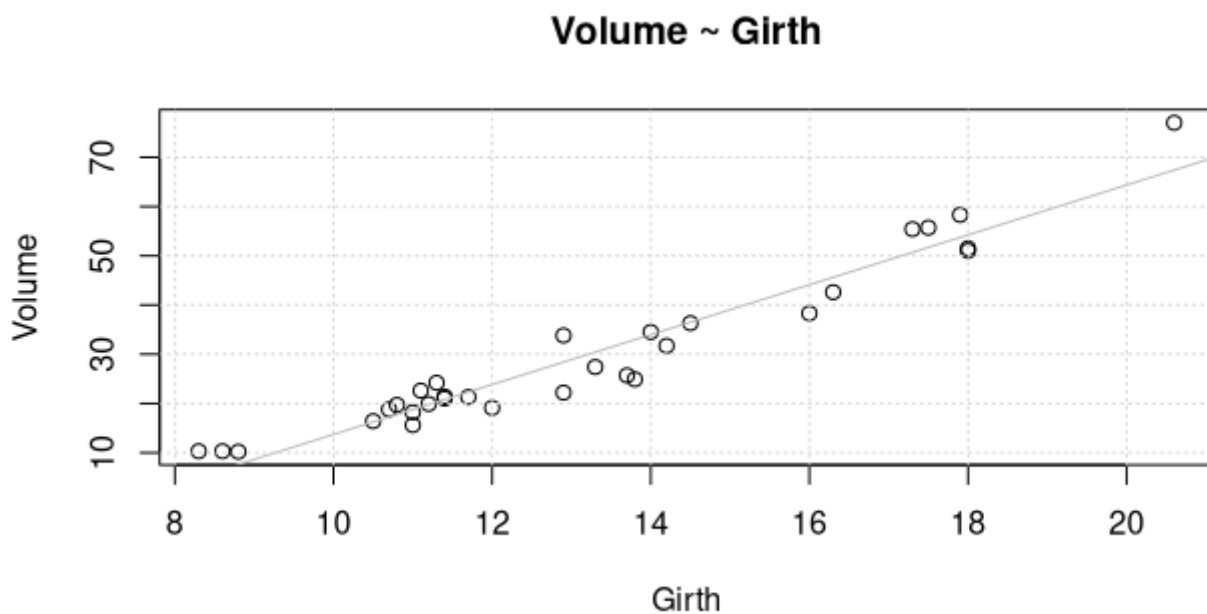
F) Quais foram os valores estimados dos coeficientes de regressão (o intercepto e a inclinação)? Como é possível interpretá-los?

```
> coefs <- gvmmodel$coefficients
> coefs
(Intercept)      Girth
   -36.943459    5.065856
```

G) Trace a reta de estimada no gráfico.

$$\hat{Y} = -36.943459 + 5.065856 \cdot X$$

```
> plot(Girth, Volume)
> abline(gvmodel, col="gray")
> grid()
> title('Volume ~ Girth')
```



H) Qual o teste que utilizamos para verificar se a relação de regressão é significativa? Faça uma conclusão sobre o teste utilizando o respectivo p-valor.

I) Qual a estatística utilizada para verificarmos se o modelo tem um alto poder explicativo para a variável resposta? Qual o valor desta estatística neste ajuste? Interprete este valor.

J) Qual o volume médio previsto para uma árvore com o diâmetro de 13 polegadas?

```
> predict(gvmodel,newdata=data.frame(Girth=c(13)),  
+         interval="prediction")  
      fit      lwr      upr  
1 28.91267 20.07634 37.74901
```