

Matrícula: 11811ETE001

Estudante: Lesly Viviane Montúfar Berrios

Disciplina: Estatística e Probabilidade - 2º período

1.

- (a) Abaixo estão dispostos os diagramas de dispersão. Ressalta-se que devido à natureza quantitativa das variáveis, o ideal é calcular a correlação(x,y) (Veja item 1.b), para obter o grau de associação linear, que será $[100 \cdot \text{correlação}(x,y)]\%$.

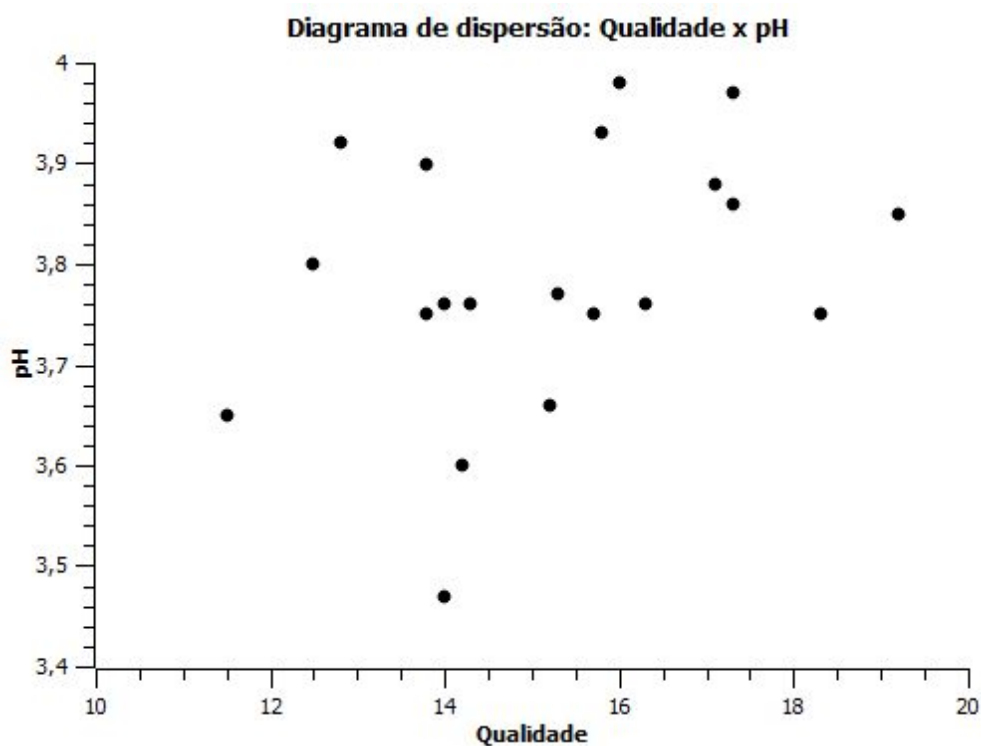


Figura 1. Figura com o diagrama de dispersão para as variáveis *Qualidade* e *pH*.

O diagrama de dispersão da Figura 1 demonstra que não há uma correlação positiva entre a *Qualidade* e o *pH*, uma vez que à medida que *Qualidade* aumenta, o *pH* aumenta.

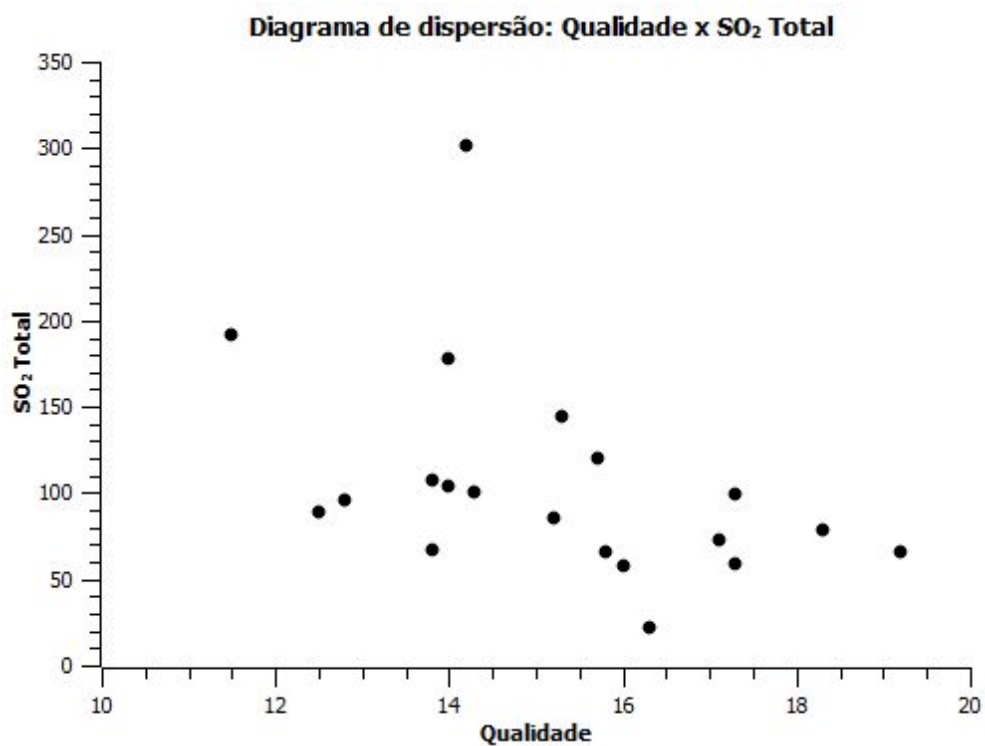


Figura 2. Figura com o diagrama de dispersão para as variáveis *Qualidade* e *SO₂ Total*.

Pelo diagrama de dispersão da Figura 2 observa-se que existe uma correlação aproximadamente linear negativa, já que à medida que a *Qualidade* aumenta, *SO₂ Total* diminui.

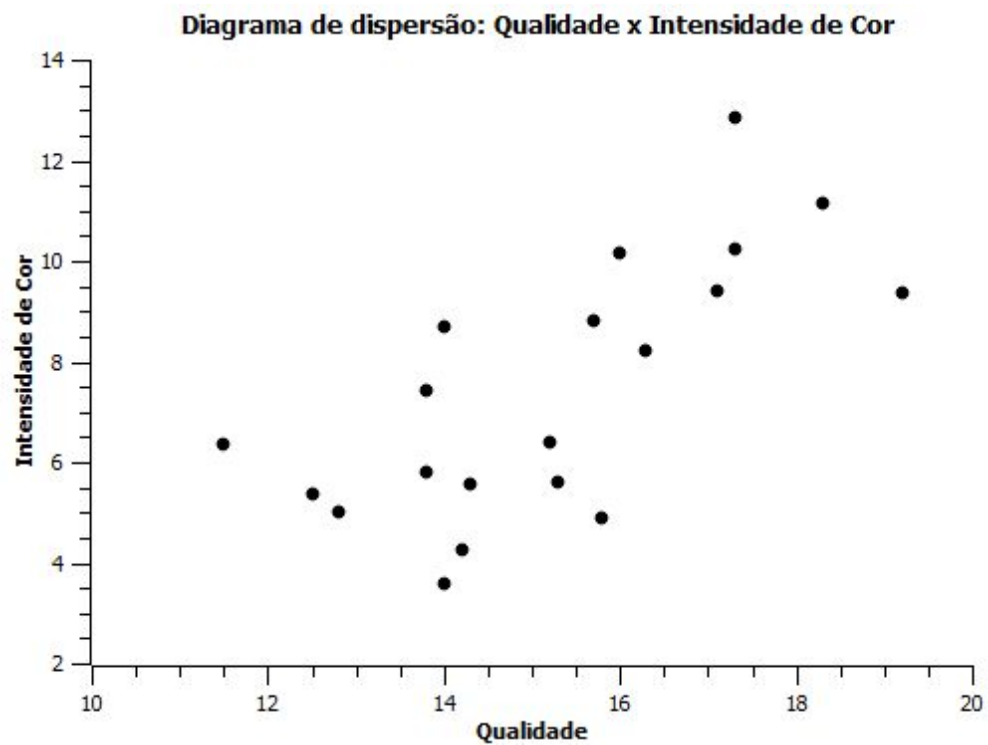


Figura 3. Figura com o diagrama de dispersão para as variáveis *Qualidade* e *Intensidade de Cor*.

Do diagrama de dispersão *Qualidade x Intensidade de Cor* (Figura 3) tem-se que à medida que a *Qualidade* aumenta a *Intensidade de Cor* também aumenta, logo existe uma correlação positiva entre essas variáveis.

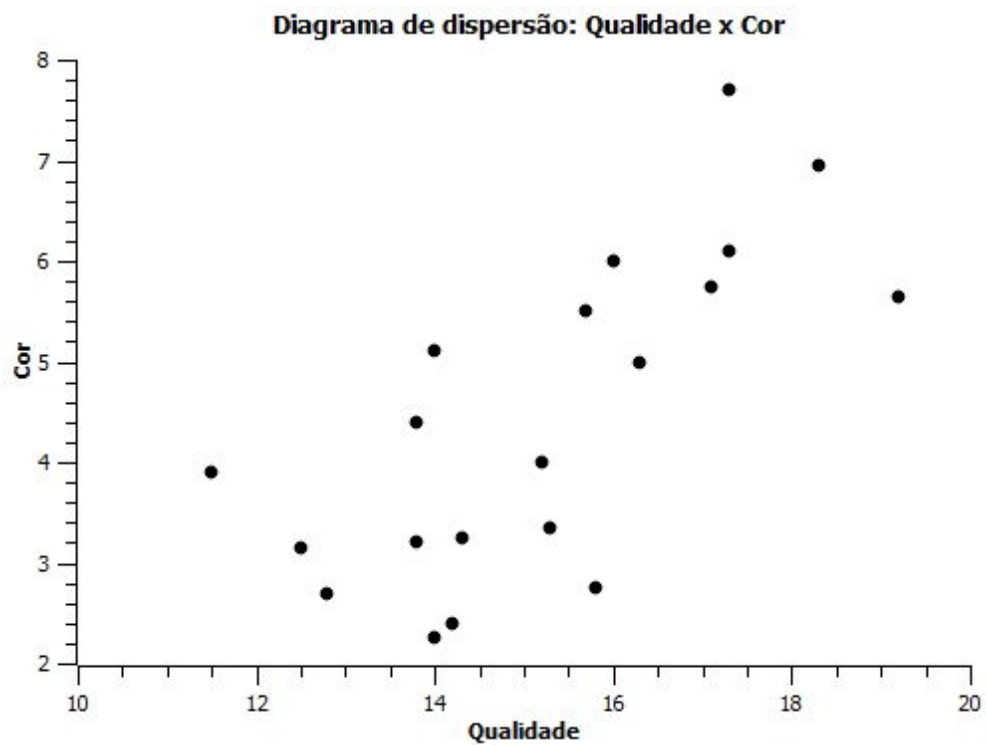


Figura 4. Figura com o diagrama de dispersão para as variáveis *Qualidade* e *Cor*.

O diagrama de dispersão *Qualidade x Cor* (Figura 4) indica que as variáveis *Qualidade* e *Cor* seguem uma correlação positiva, uma vez à medida que a *Cor* aumenta a *Qualidade* aumenta.

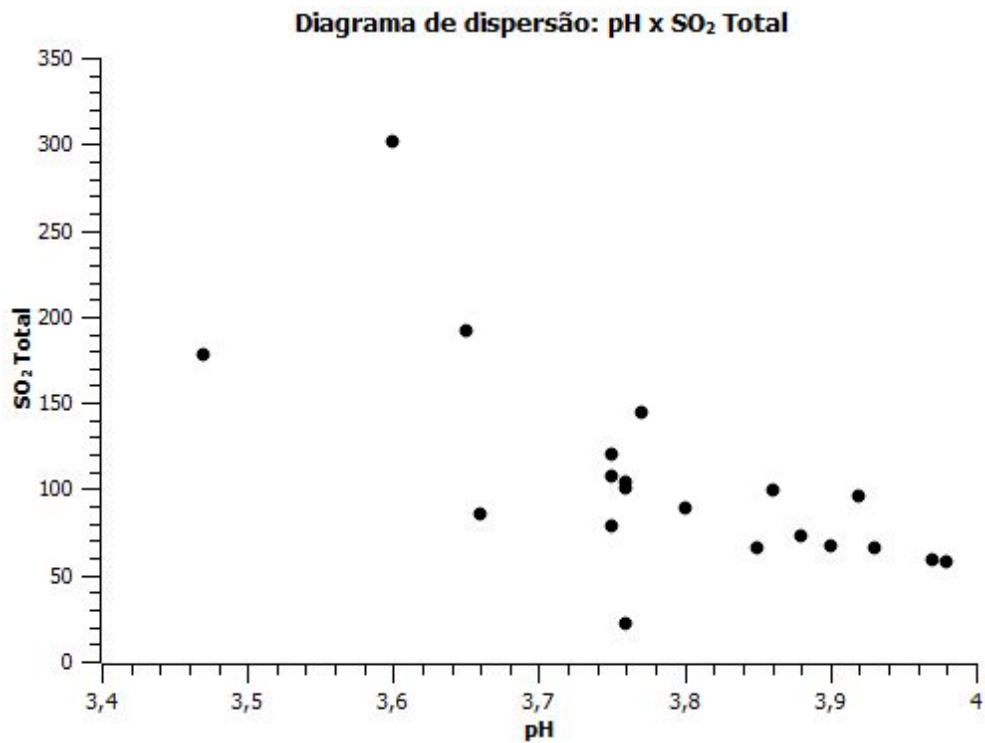


Figura 5. Figura com o diagrama de dispersão para as variáveis pH e SO_2 Total.

O diagrama de dispersão $pH \times SO_2$ Total (Figura 5) indica que as variáveis pH e SO_2 Total seguem uma correlação negativa, uma vez à medida que a pH aumenta a SO_2 Total diminui.

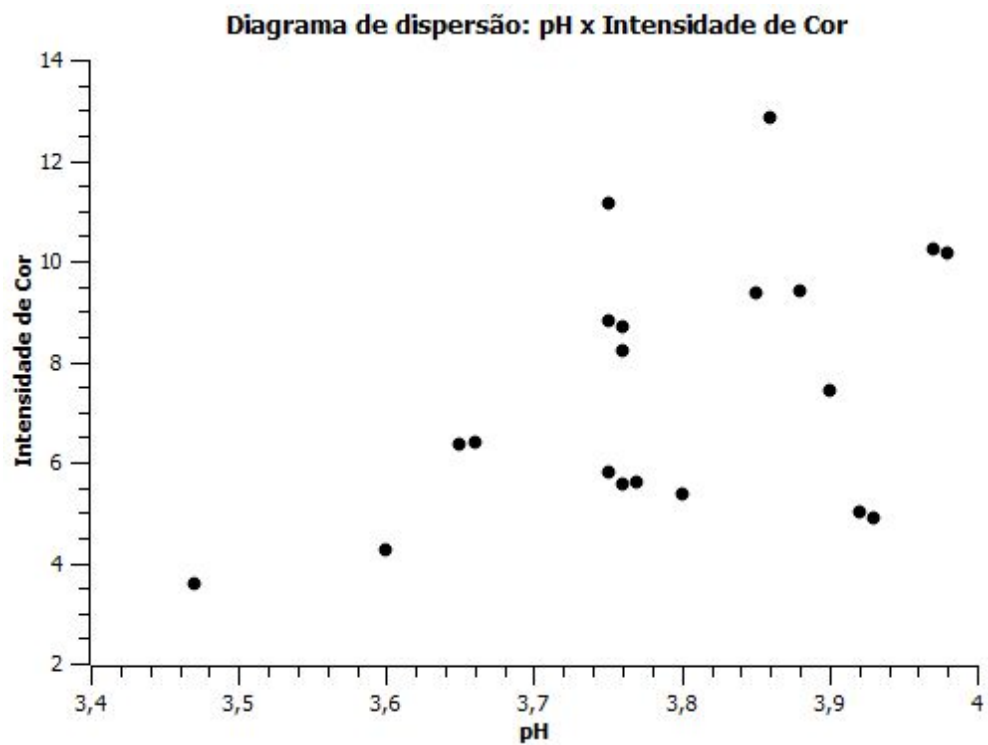


Figura 6. Figura com o diagrama de dispersão para as variáveis *pH* e *Intensidade de Cor*.

O diagrama de dispersão *pH x Intensidade de Cor* (Figura 6) indica que as variáveis *pH* e *Intensidade de Cor* seguem uma correlação positiva, uma vez à medida que a *pH* aumenta a *Intensidade de Cor* aumenta.

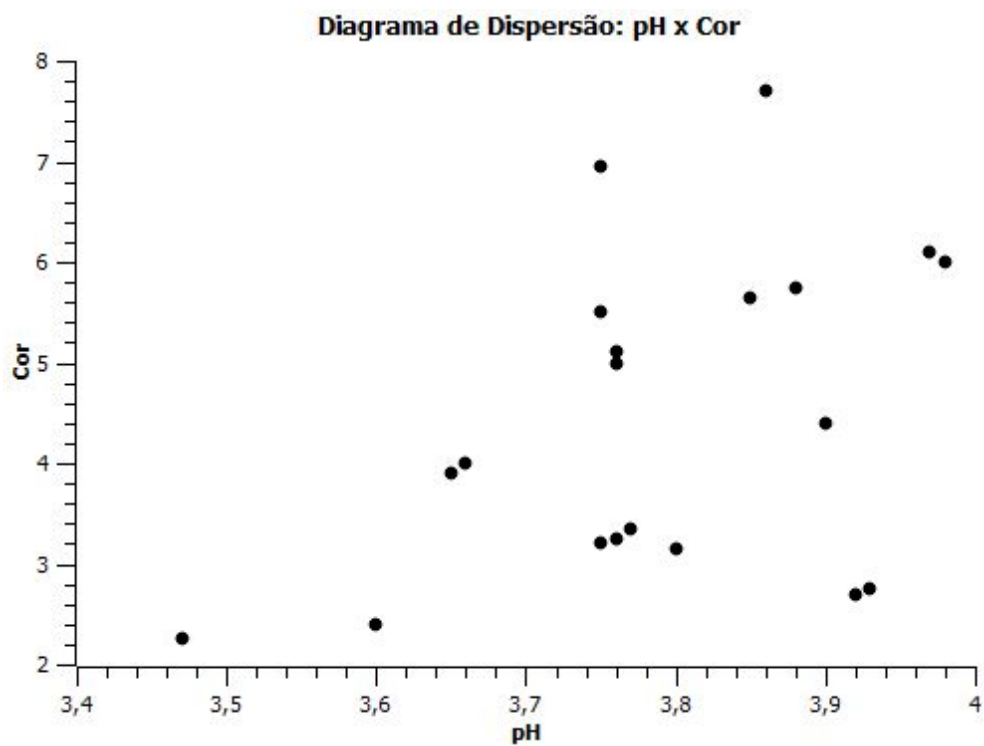


Figura 7. Figura com o diagrama de dispersão para as variáveis *pH* e *Cor*.

O diagrama de dispersão *pH x Cor* (Figura 7) indica que as variáveis *pH* e *Cor* seguem uma correlação positiva, uma vez à medida que a *pH* aumenta a *Cor* aumenta.

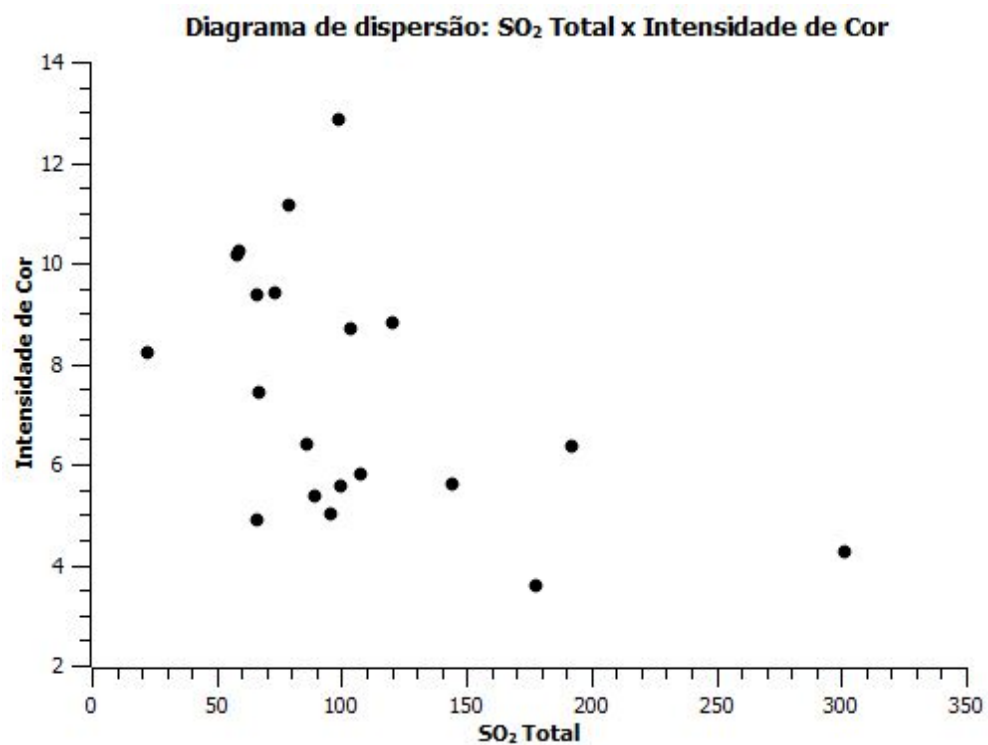


Figura 8. Figura com o diagrama de dispersão para as variáveis SO_2 Total e Intensidade de Cor.

O diagrama de dispersão SO_2 Total x Intensidade de Cor (Figura 8) indica que as variáveis SO_2 Total e Intensidade de Cor seguem uma correlação negativa, uma vez à medida que a SO_2 Total aumenta a Intensidade de Cor diminui.

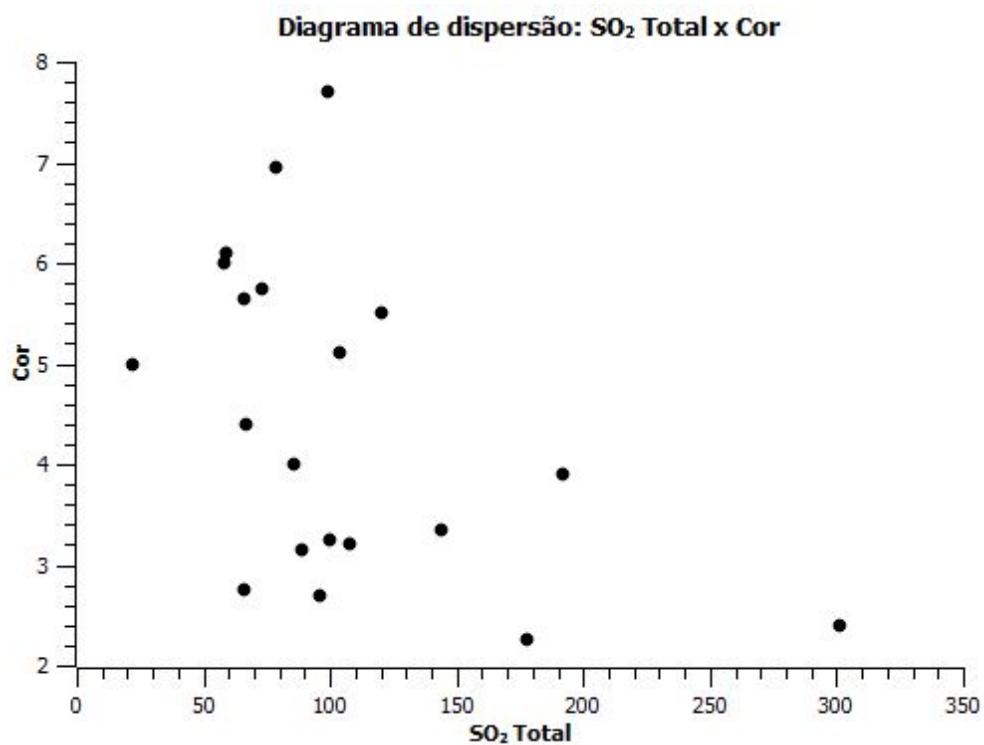


Figura 9. Figura com o diagrama de dispersão para as variáveis SO_2 Total e Cor.

O diagrama de dispersão SO_2 Total x Cor (Figura 9) indica que as variáveis SO_2 Total e Cor seguem uma correlação negativa, uma vez à medida que a SO_2 Total aumenta a Cor diminui.

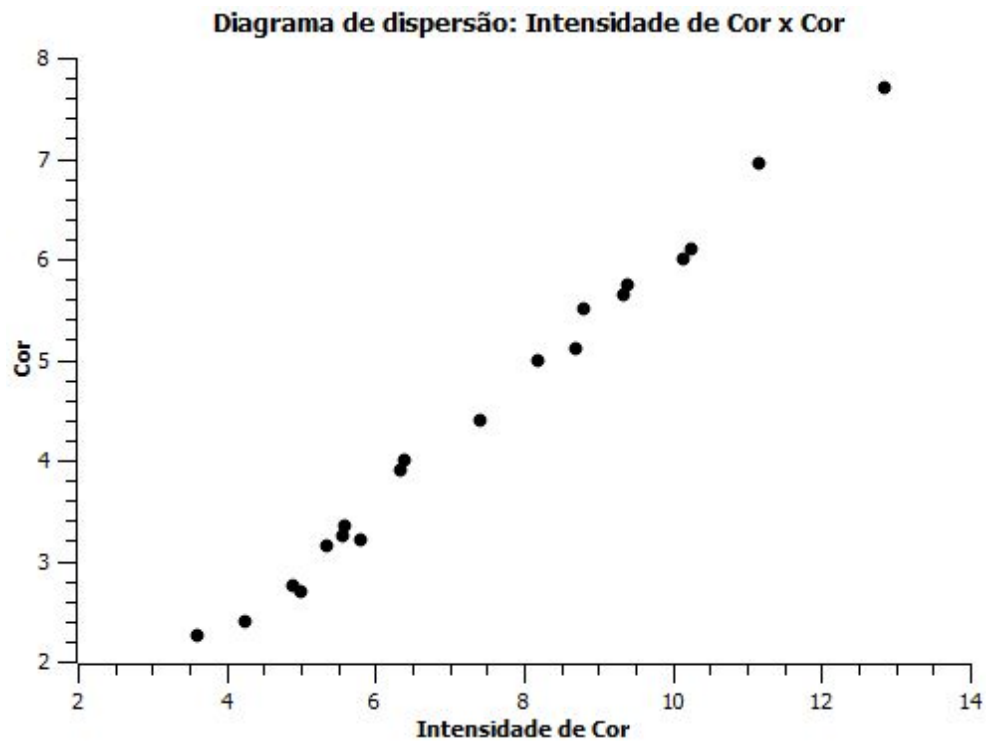


Figura 10. Figura com o diagrama de dispersão para as variáveis *Intensidade de Cor* e *Cor*.

O diagrama de dispersão *Intensidade de Cor x Cor* indica que as variáveis *Intensidade de Cor* e *Cor* seguem uma correlação positiva, uma vez à medida que a *Intensidade de Cor* aumenta a *Cor* aumenta.

(b) O cálculo do coeficiente correlação linear para o cruzamento de todas as variáveis foi realizado via-Excel e o resultado está disposto na Tabela 1, com aproximação de 4 casas decimais.

Tabela 1: Tabela com os valores de correlação linear fornecidos pelo Excel, por meio do comando =CORREL().

	<i>Qualidade</i>	<i>pH</i>	<i>SO₂ total</i>	<i>Int. de cor</i>	<i>Cor</i>
<i>Qualidade</i>	1	0,3492	-0,4447	0,7024	0,7121
<i>pH</i>	0,3492	1	-0,6786	0,4819	0,4304
<i>SO₂ total</i>	-0,4447	-0,6786	1	-0,4922	-0,4797
<i>Int. de cor</i>	0,7024	0,4819	-0,4922	1	0,9962
<i>Cor</i>	0,7121	0,4304	-0,4797	0,9962	1

(c) Análise das correlações calculadas e explicação dos resultados obtidos.

As correlações calculadas entre as variáveis *Cor* e *Intensidade de Cor* são as maiores (em módulo), o que significa que possuem um alto grau de associação linear, ou seja se aproxima mais de uma reta, que facilmente verificado na Figura 10. Analogamente, as correlações calculadas entre as variáveis *Qualidade* e *Int. de Cor*, e *Qualidade* e *Cor* são pouco menores e significam que se aproximam menos de uma reta em relação às variáveis da Figura 10, enquanto que a correlação entre as variáveis *Qualidade* e *pH* é baixa, já que o gráfico da Figura 1 mostra valores muito dispersos. Ademais, a reta de regressão linear para os casos mencionados será crescente, uma vez que a correlação é positiva.

Para as variáveis de correlação negativa, o raciocínio é o mesmo, com a diferença que quanto maior o módulo da correlação, mais o gráfico de dispersão se aproxima de uma reta decrescente. Por exemplo, entre as correlações negativas, as correlações entre as variáveis *pH* e *SO₂ total* possuem o maior módulo, portanto o gráfico entre essas variáveis (Veja Figura 5) se aproxima mais de uma reta decrescente comparada aos gráficos entre as variáveis *Qualidade* e *SO₂ total* (Figura 2), *SO₂ total* e *Int. de Cor* (Figura 8) e *SO₂ total* e *Cor* (Figura 9).

(d) O histograma e boxplot para cada variável é apresentado a seguir, juntamente com o comentário sobre a simetria ou assimetria delas.

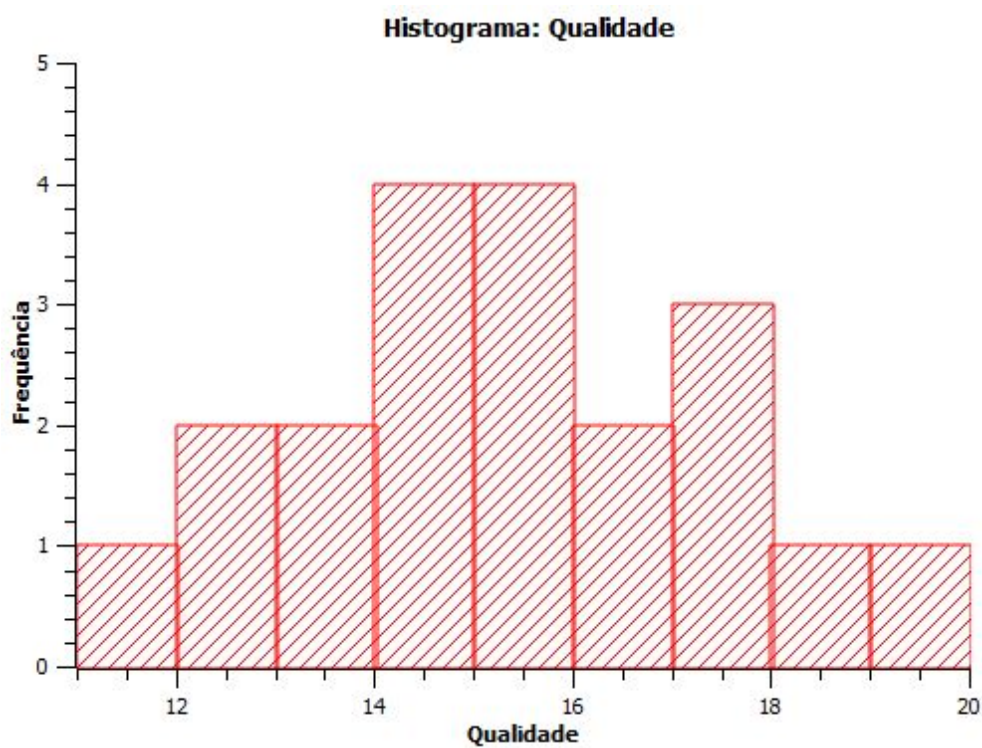


Gráfico 1. Histograma dos dados de *Qualidade*.

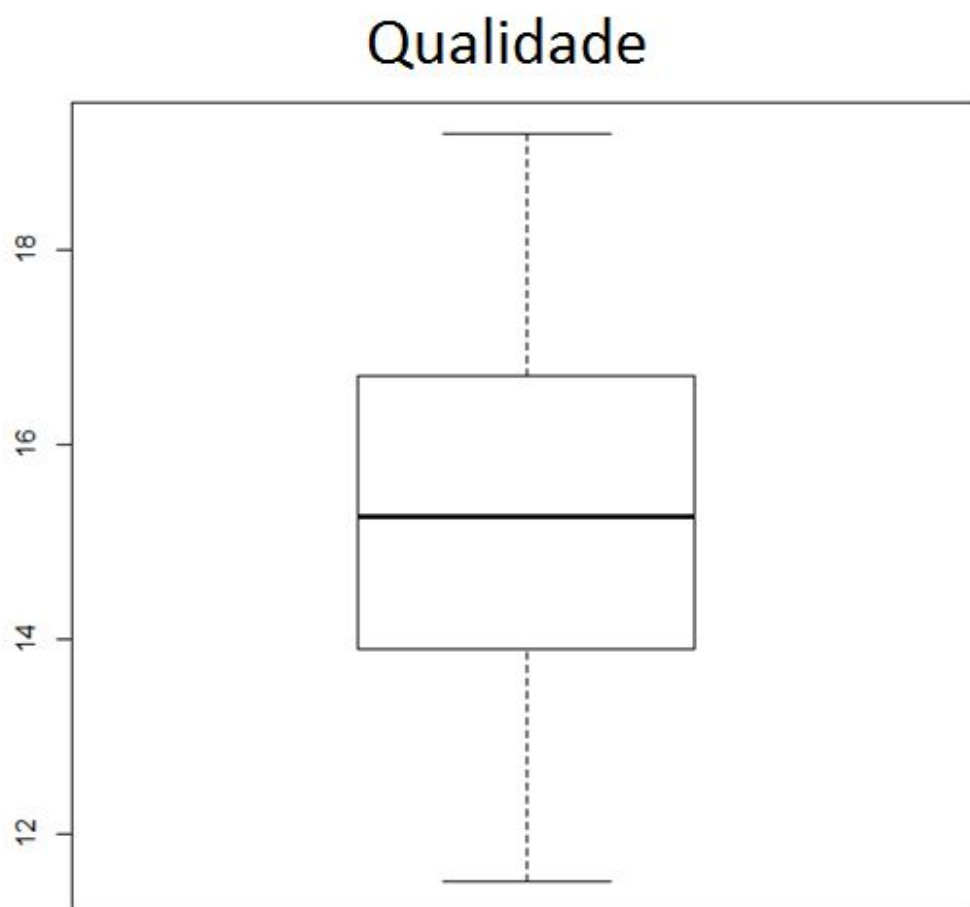


Gráfico 2. Boxplot dos dados de *Qualidade*.

A partir dos gráficos 1 e 2 conclui-se que a variável *Qualidade* é razoavelmente simétrica, já que, no Gráfico 3, observa-se que os dados se aproximam da curva de distribuição normal, mesmo com poucos dados experimentais.

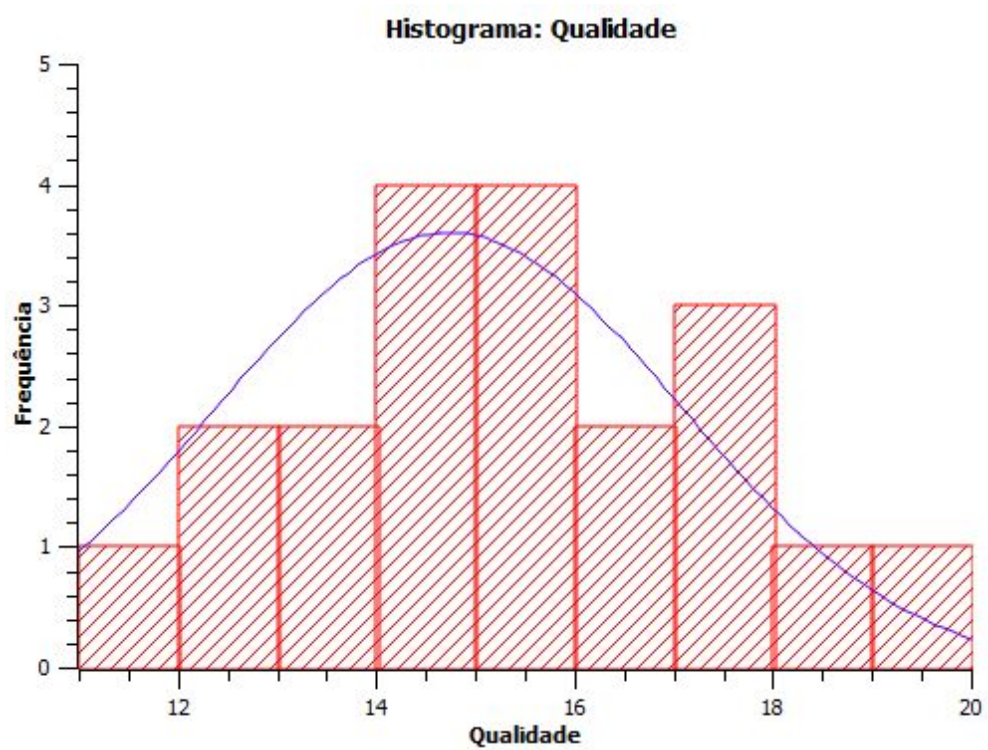


Gráfico 3. Histograma para os dados de *Qualidade* com inclusão da curva de distribuição.

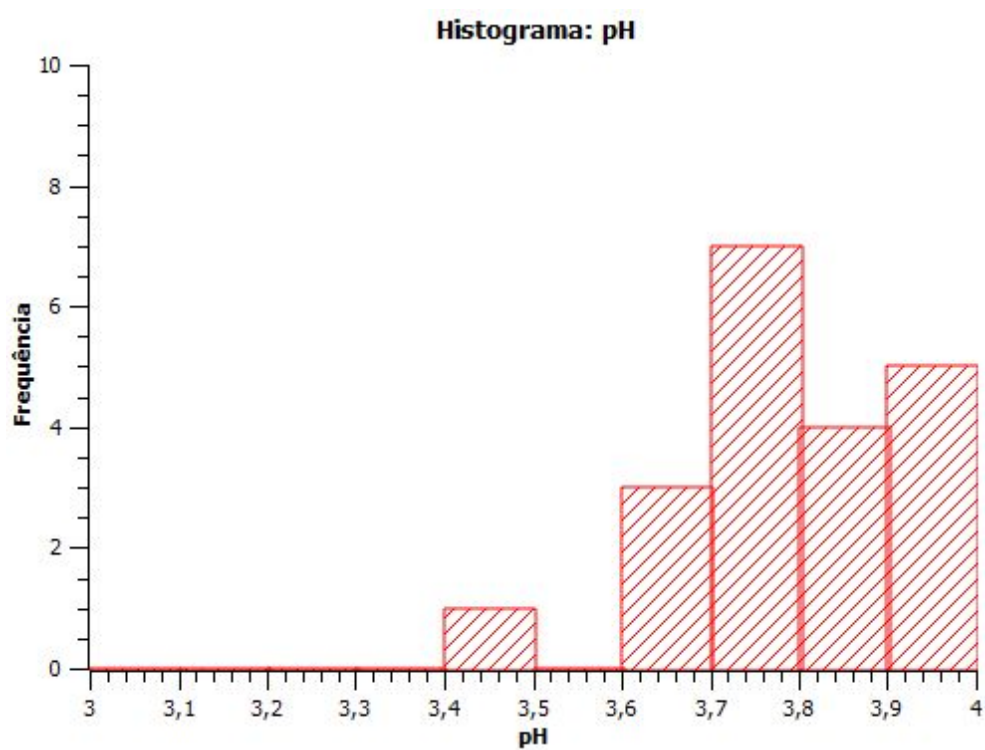


Gráfico 4. Histograma para os dados de *pH*.

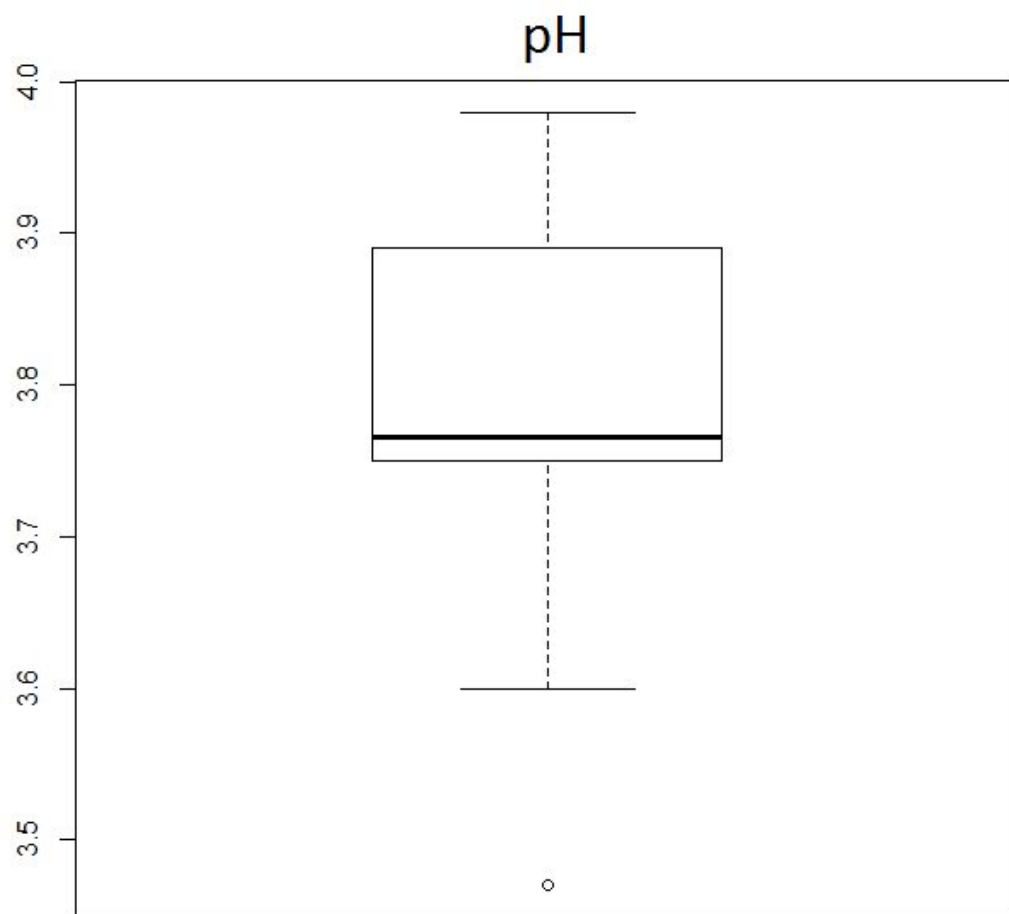


Gráfico 5. Boxplot dos dados de pH .

A partir dos gráficos 4 e 5 conclui-se que a variável pH é razoavelmente simétrica, já que, no Gráfico 6, observa-se que os dados se aproximam da curva de distribuição normal, mesmo com poucos dados experimentais.

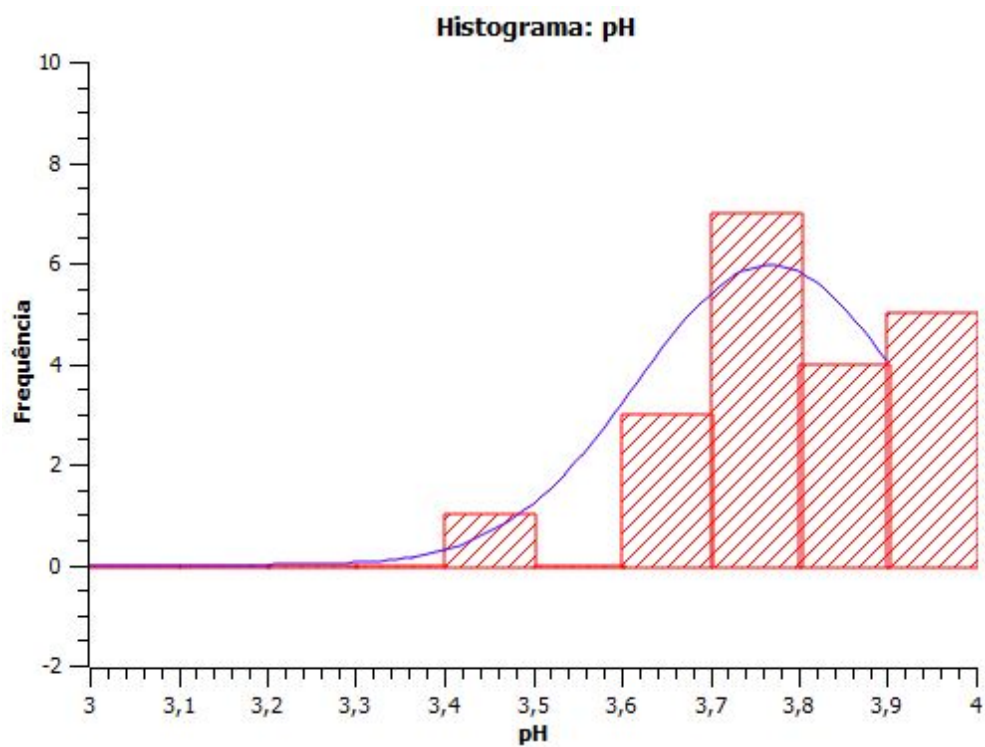


Gráfico 6. Histograma para os dados de *pH* com inclusão da curva de distribuição.

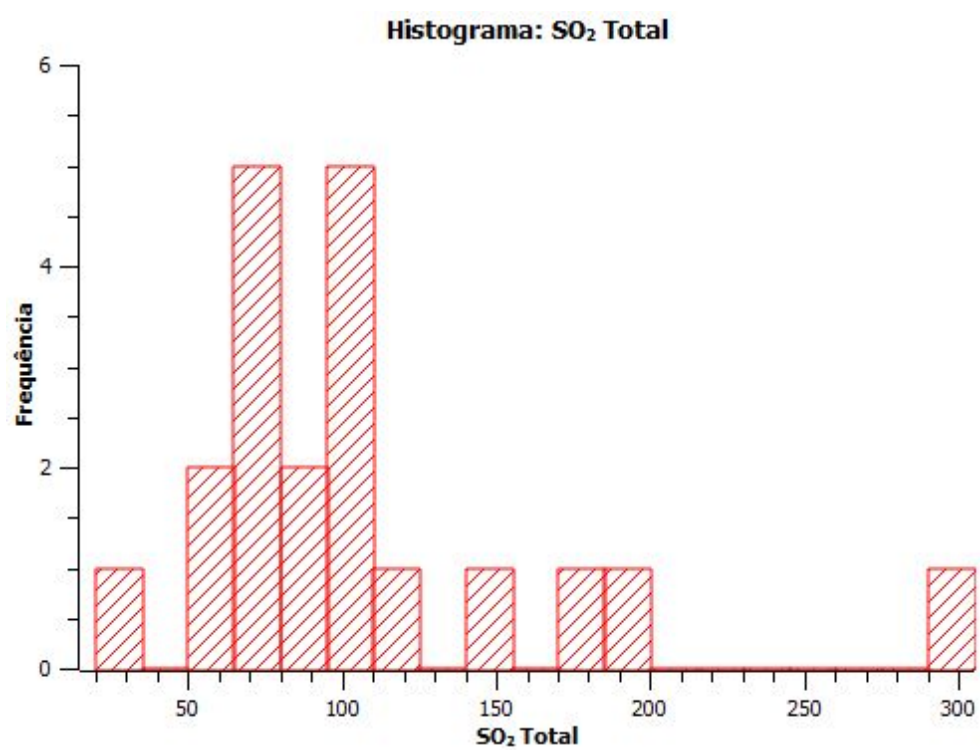


Gráfico 7. Histograma para os dados de *SO₂ Total*.

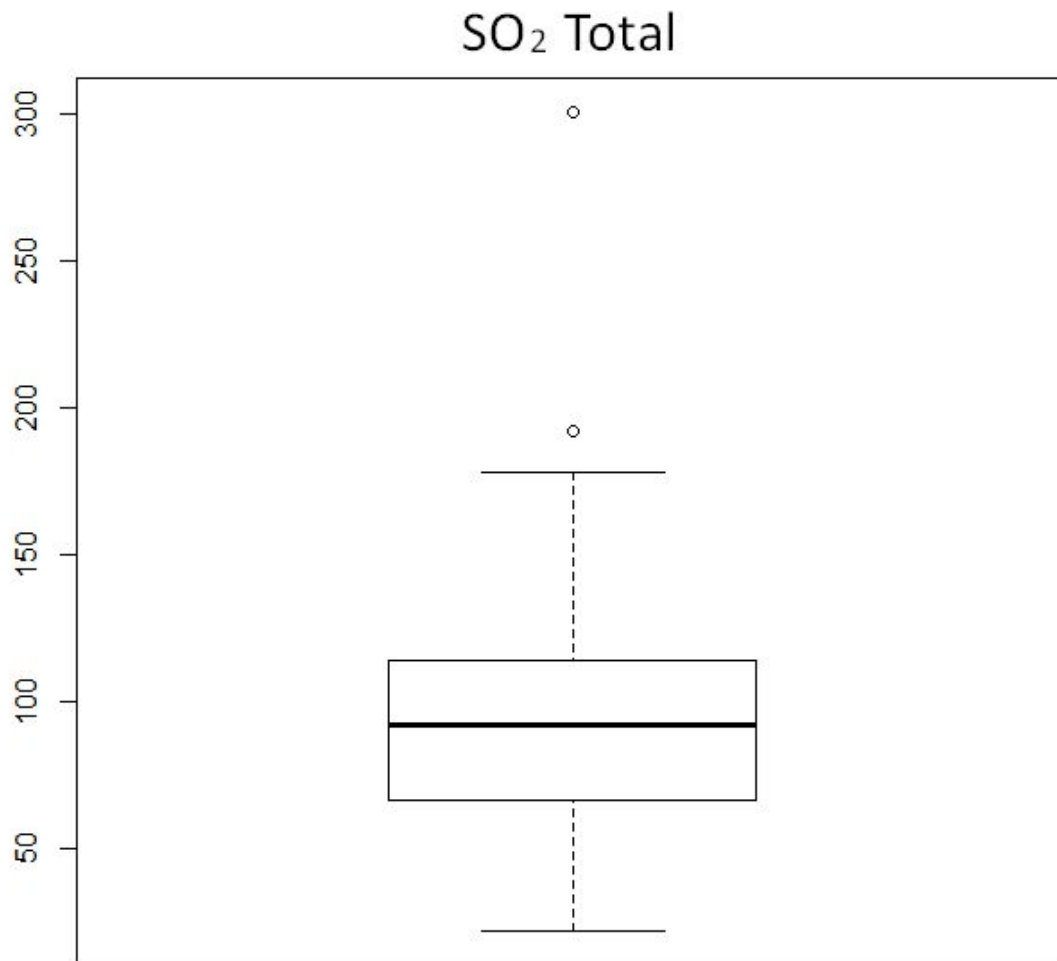


Gráfico 8. Boxplot dos dados de *SO₂ total* .

Os dados de *SO₂ total* são muito assimétricos (Veja Gráfico 7 e 8), uma vez que no histograma percebe-se uma grande diferença em relação ao gráfico de distribuição normal, sendo impossível traçar uma curva aproximadamente gaussiana. Além disso, pelo boxplot do Gráfico 8 percebe-se ainda a existência de 2 outliers.

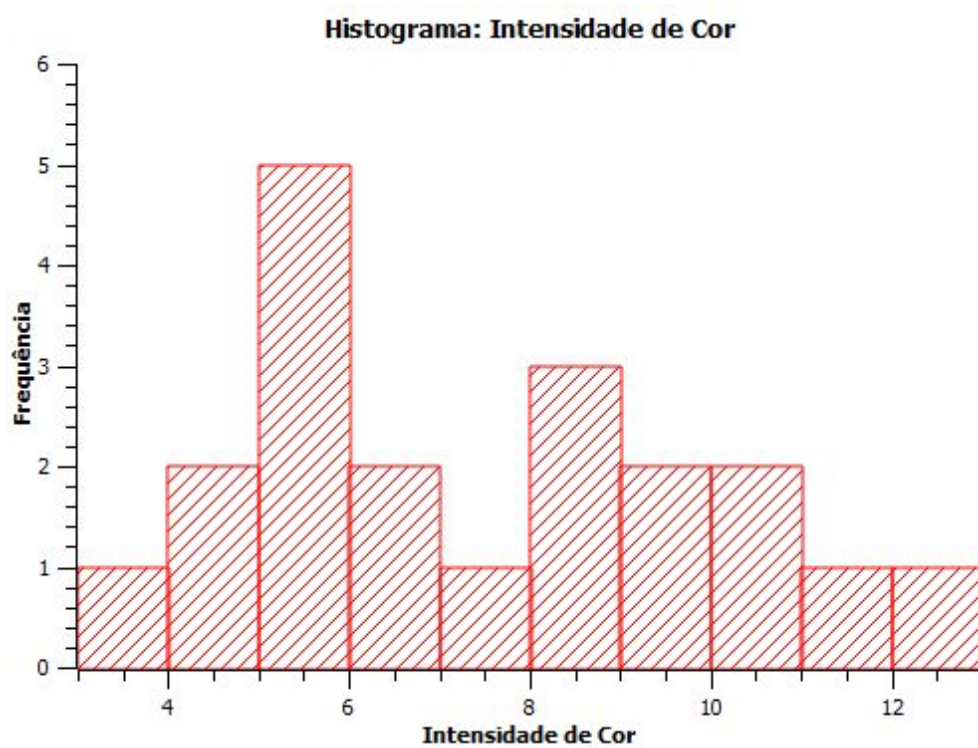


Gráfico 9. Histograma para os dados de *Intensidade de Cor*.

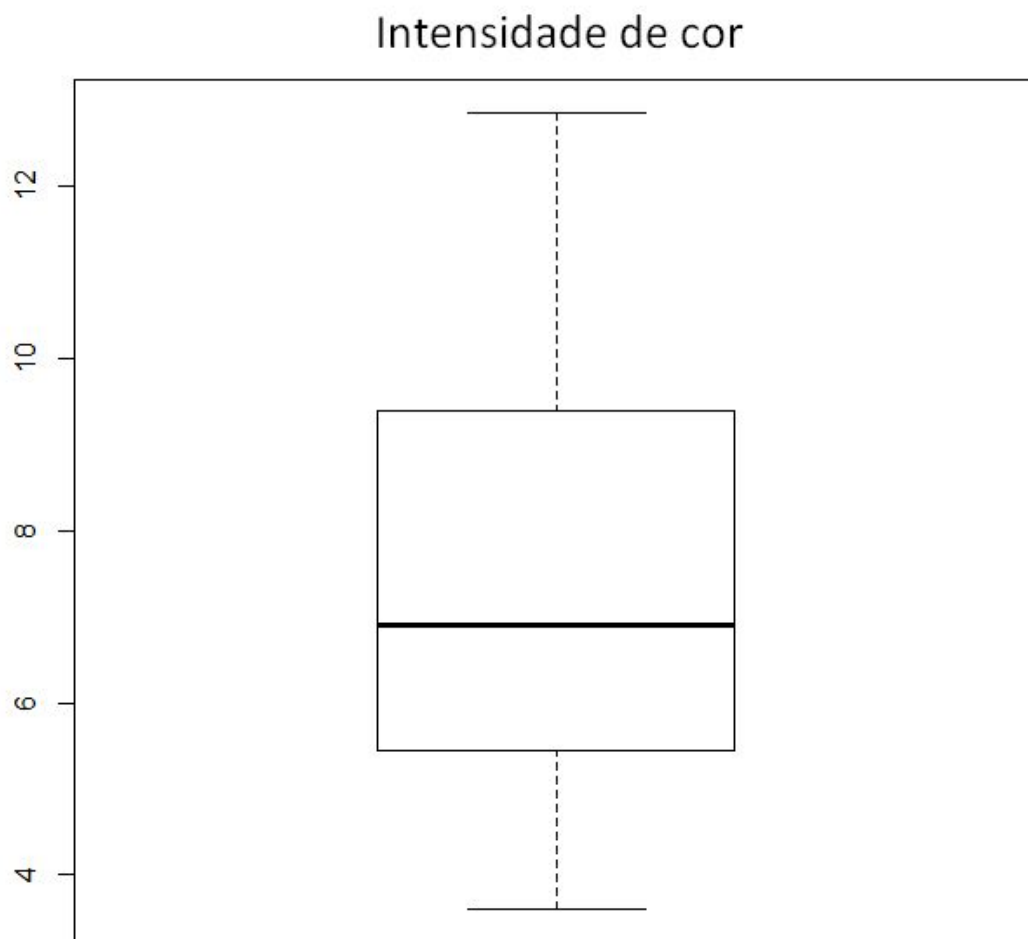


Gráfico 10. Boxplot dos dados de Intensidade de Cor

A partir dos gráficos 9 e 10 conclui-se que a variável *Intensidade de Cor* é pouco assimétrica, já que, no Gráfico 11, observa-se que os dados se aproximam da curva de distribuição normal somente para dados experimentais no intervalo [3,7].

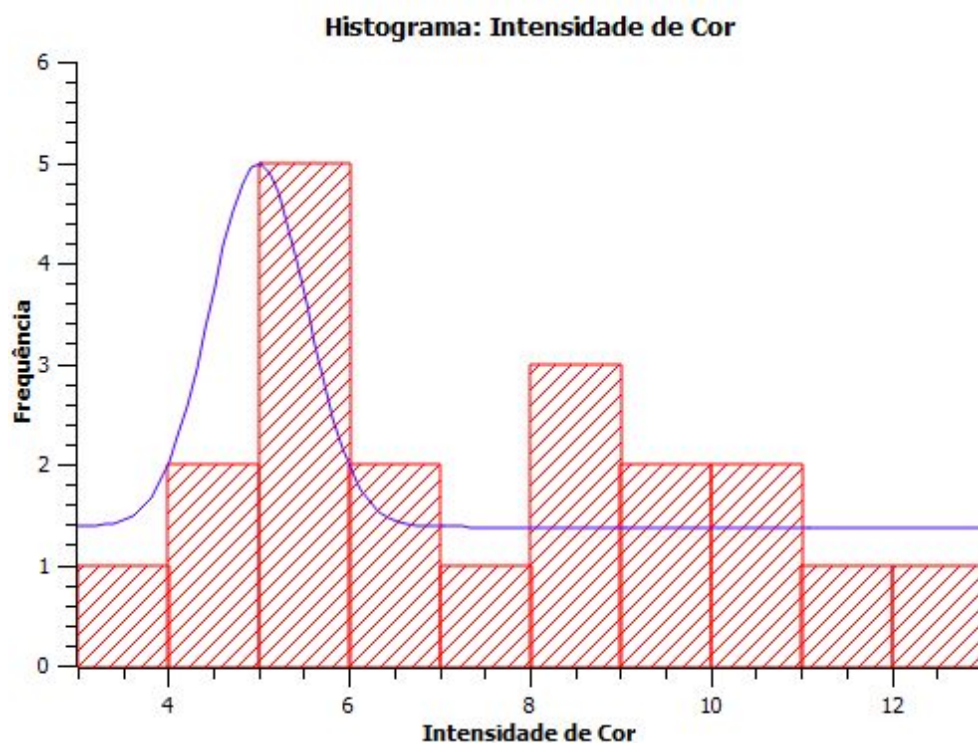


Gráfico 11. Histograma para os dados de *Intensidade de Cor* com inclusão da curva de distribuição.

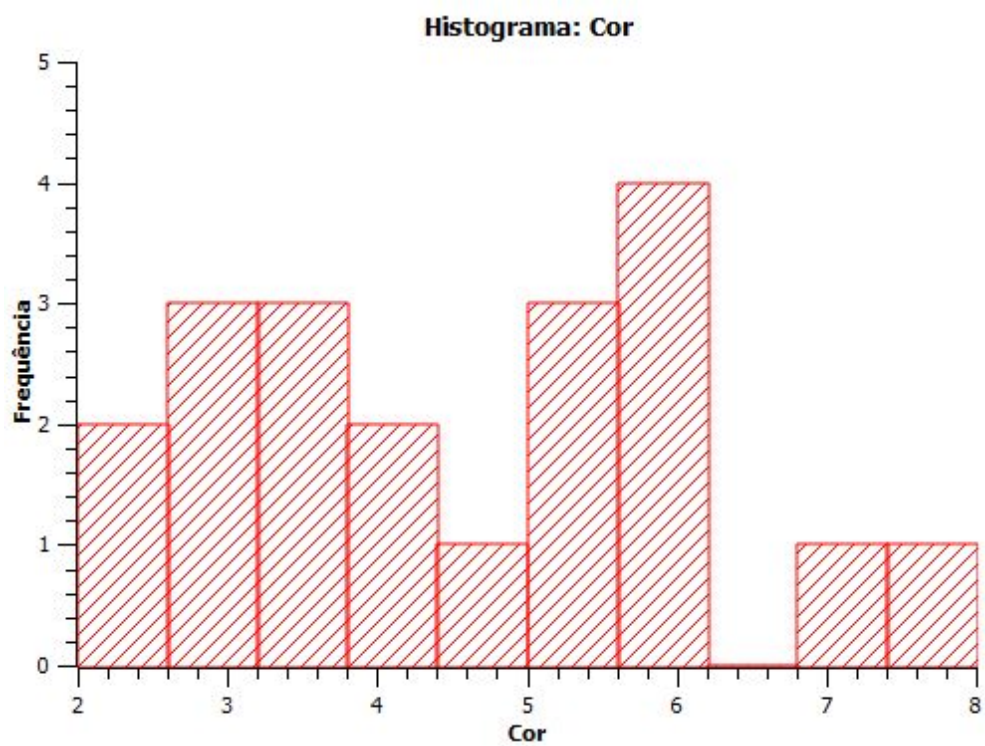


Gráfico 12. Histograma para os dados de *Cor*.

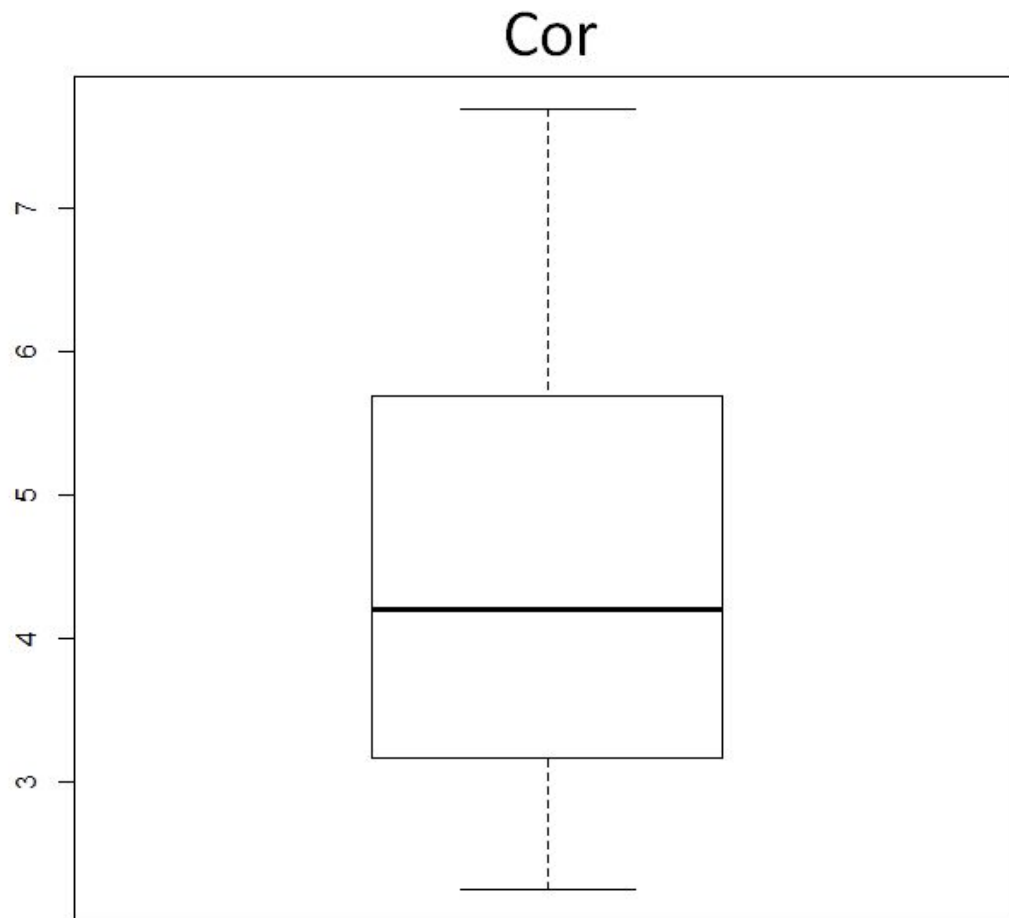


Gráfico 13. Boxplot dos dados de *Cor*.

A partir dos gráficos 12 e 13 conclui-se que a variável *Cor* é razoavelmente assimétrica, já que, no Gráfico 14, observa-se que os dados somente se aproximam da curva de distribuição normal no intervalo [2,5].

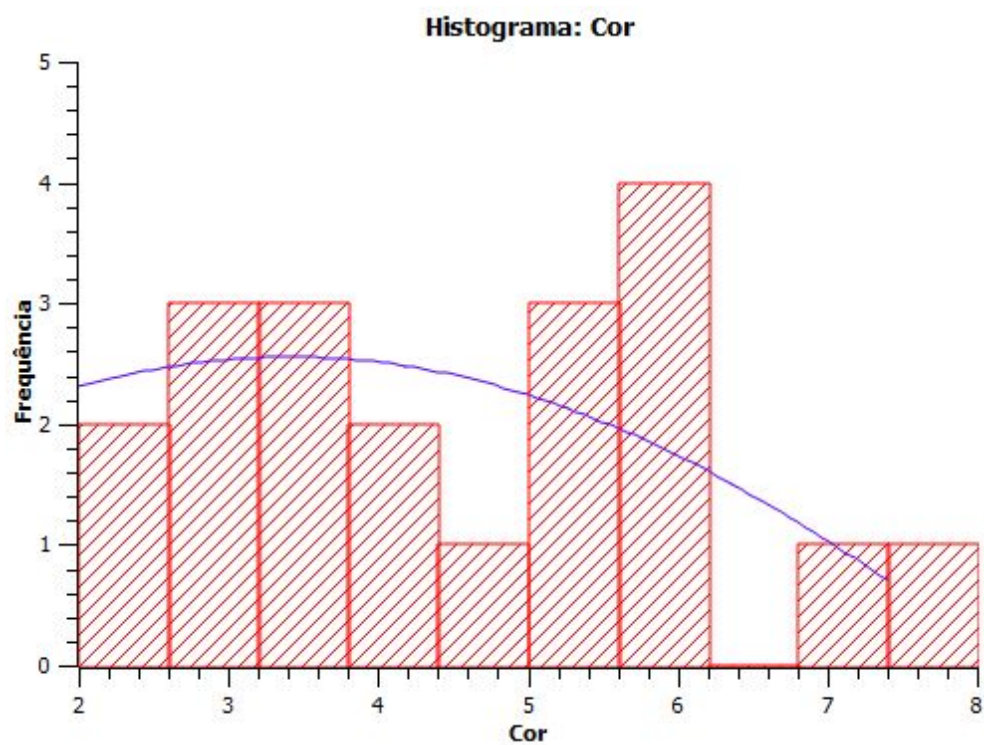


Gráfico 14. Histograma para os dados de *Cor* com inclusão da curva de distribuição.

