

**INSTITUTO FEDERAL DE EDUCAÇÃO
CIÊNCIA E TECNOLOGIA DE MINAS GERAIS -
CAMPUS BAMBUÍ**

**BACHARELADO EM ENGENHARIA DE
COMPUTAÇÃO**

GABRIELA DÂMASO REZENDE

TÓPICOS ESPECIAIS EM ENGENHARIA DE COMPUTAÇÃO

RELATÓRIO 2

BAMBUÍ - MG

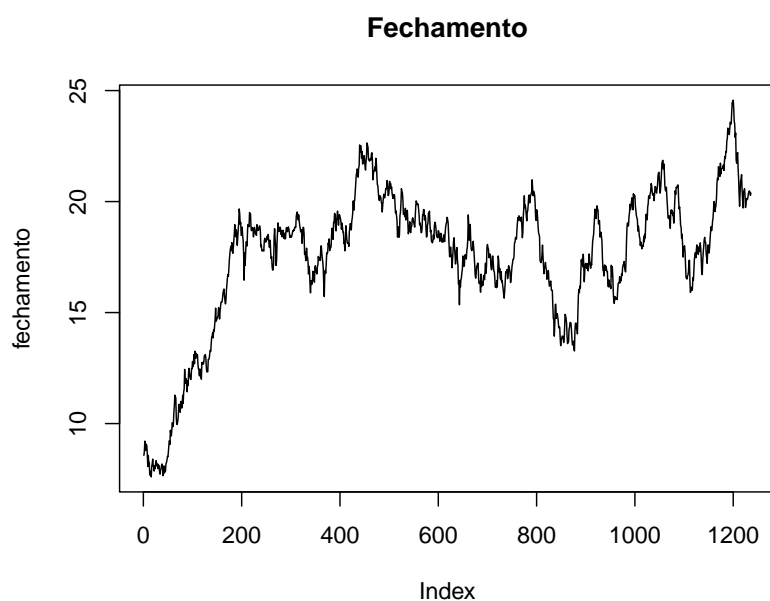
2023

1 Questão 1

1.1 Análise de gráfico

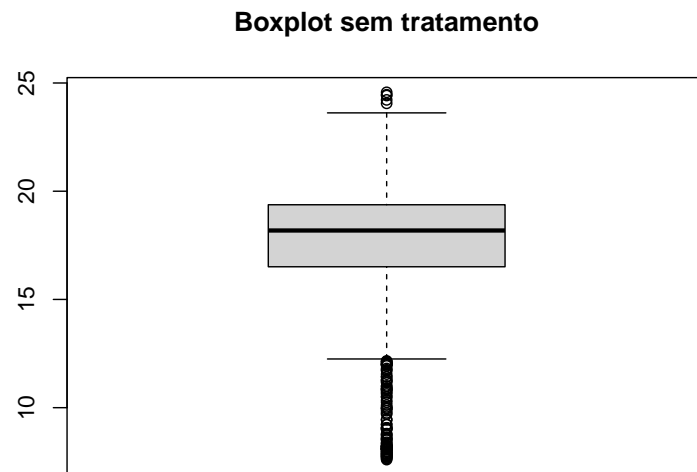
As imagens a seguir são os primeiros gráficos gerados com base na coluna fec (Fechamento). Portanto, não apresentam tratamento e se é possível notar a presença de *outliers*.

Figura 1: Gráfico 1



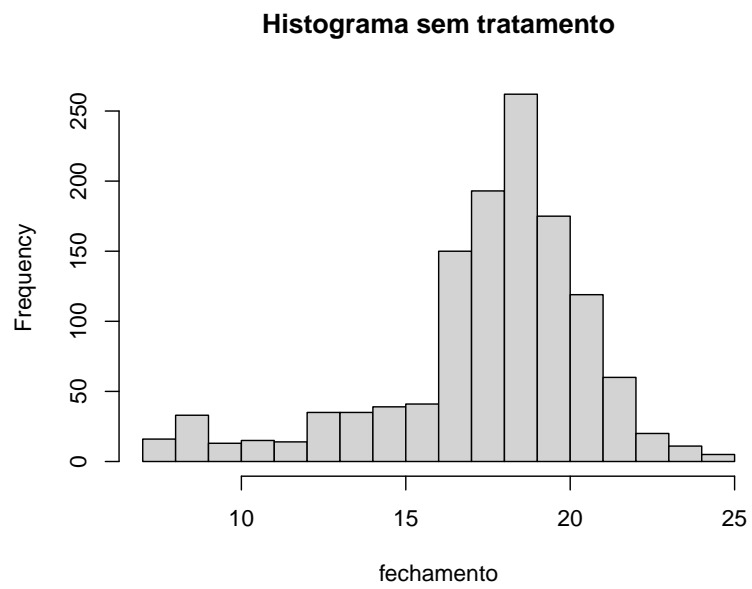
Fonte: O autor.

Figura 2: Gráfico 2



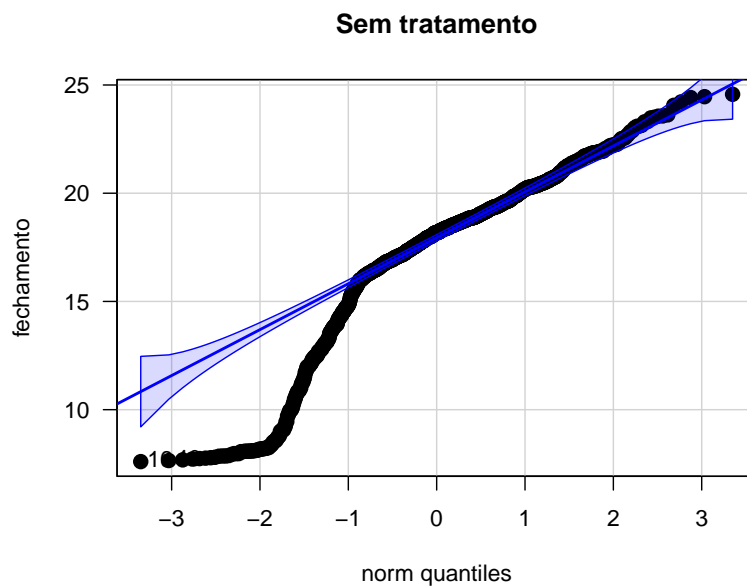
Fonte: O autor.

Figura 3: Gráfico 3



Fonte: O autor.

Figura 4: Gráfico 4



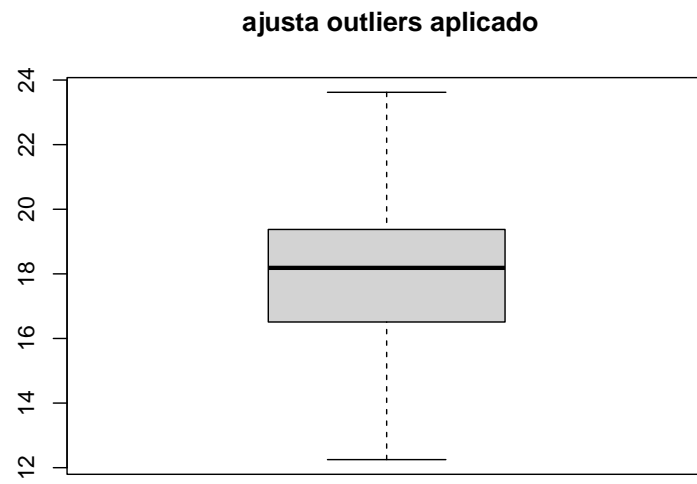
Fonte: O autor.

Na Figura 2 podemos perceber alguns possíveis *outliers* na parte superior e inferior do gráfico. Já na Figura 3 ao tomar como referência o eixo x entre os valores 0 e aproximadamente 13 os *outliers* são perceptíveis. Na Figura 4 podemos ver alguns pontos na parte inferior esquerda do gráfico e que estão abaixo da faixa azul.

1.2 Detecção

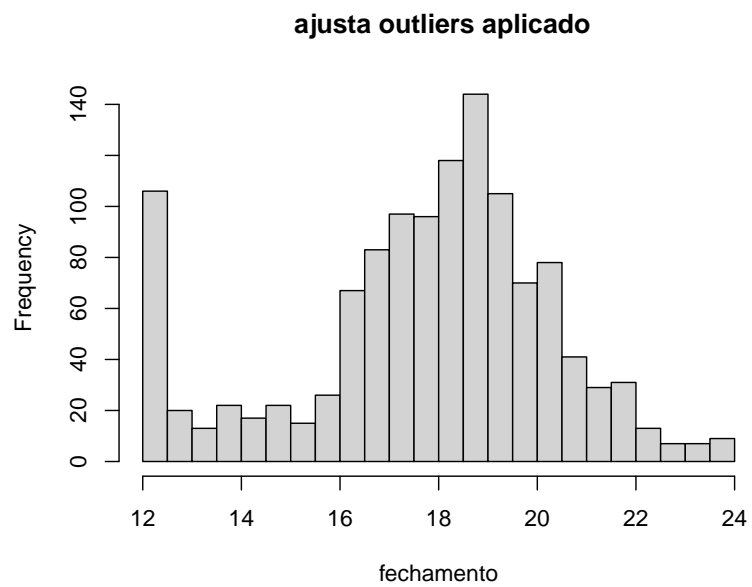
Para a realização desta etapa foi aplicado a função `ajustaOutliers` e obtivemos o seguinte resultado:

Figura 5: Gráfico 5



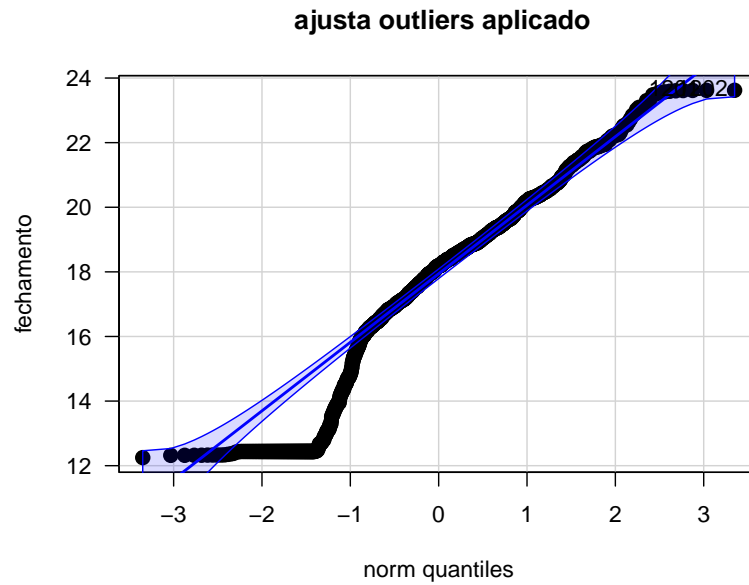
Fonte: O autor.

Figura 6: Gráfico 6



Fonte: O autor.

Figura 7: Gráfico 7



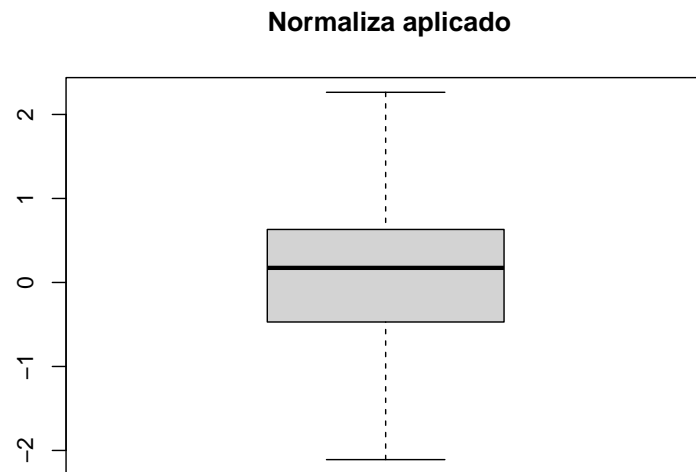
Fonte: O autor.

Pode-se notar que na Figura 5 onde podíamos ver claramente a presença de *outliers* sofreu grandes mudanças e agora os *outliers* não são mais encontrados nos locais antes observados.

1.3 Normaliza

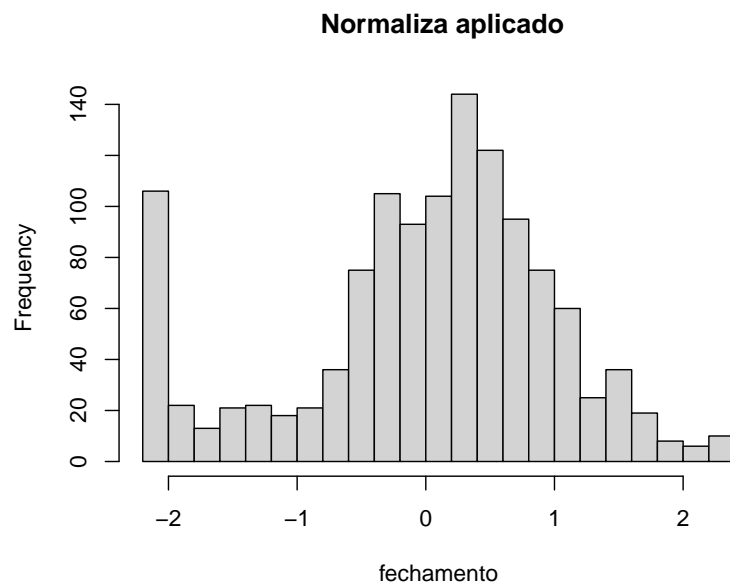
Ao aplicar a função `normaliza` foram obtidos os seguintes gráficos:

Figura 8: Gráfico 8



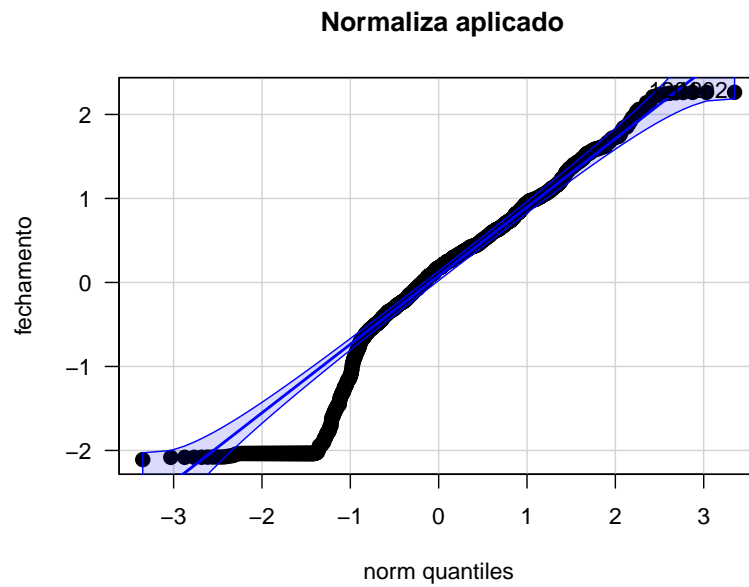
Fonte: O autor.

Figura 9: Gráfico 9



Fonte: O autor.

Figura 10: Gráfico 10



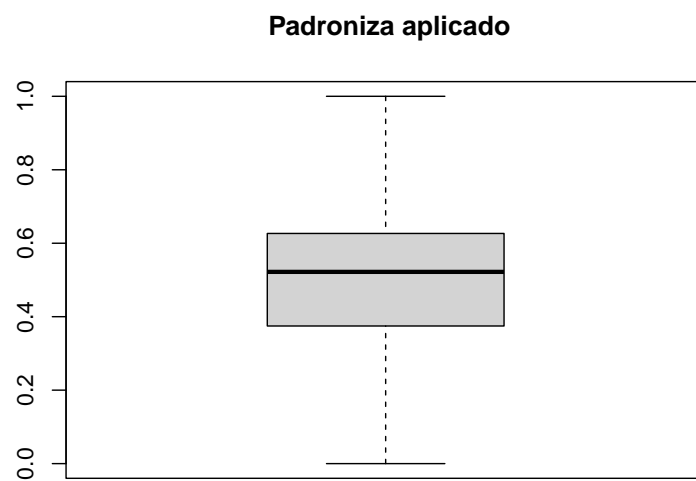
Fonte: O autor.

Com a normalização feita, pode-se notar que na Figura 10 no canto inferior direito ainda existe uma parte que se encontra fora da faixa azul. Porém, percebe-se também que a faixa azul mudou sua posição e se encontra mais ao lado inferior direito, possibilitando assim, a abrangência de mais dados.

1.4 Padronização

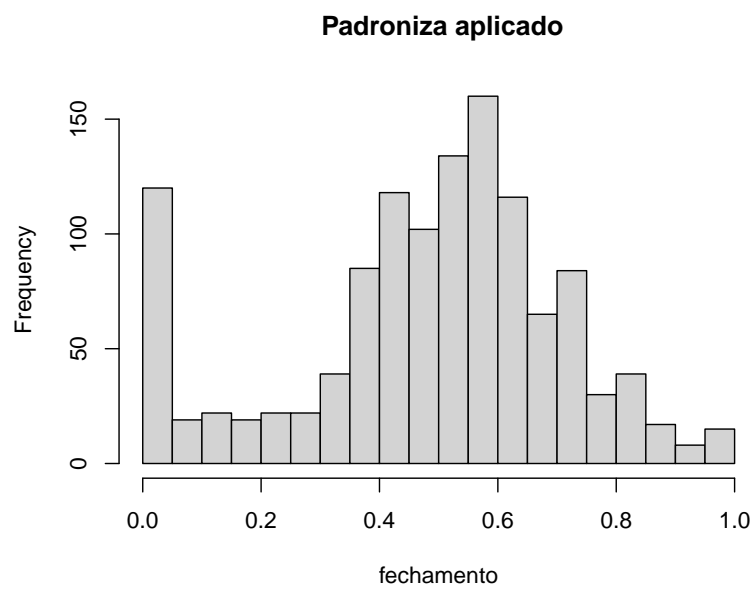
Ao aplicar a função de padronização obtêm-se os seguintes gráficos:

Figura 11: Gráfico 11



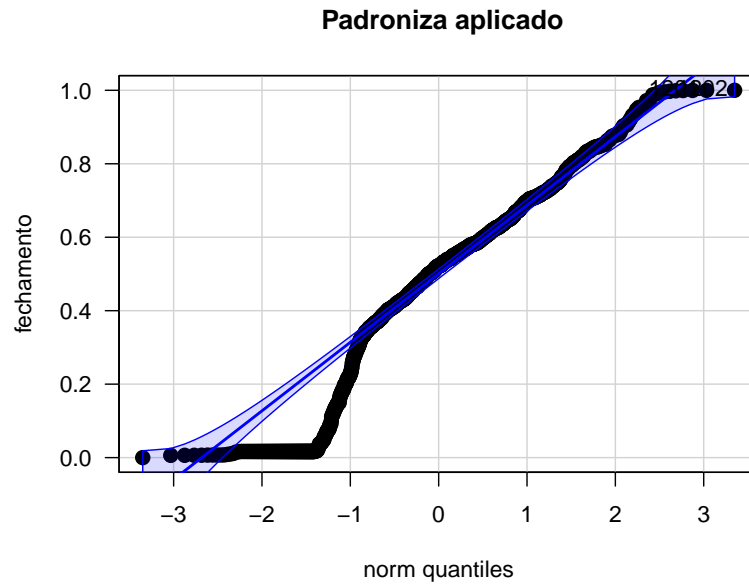
Fonte: O autor.

Figura 12: Gráfico 12



Fonte: O autor.

Figura 13: Gráfico 13



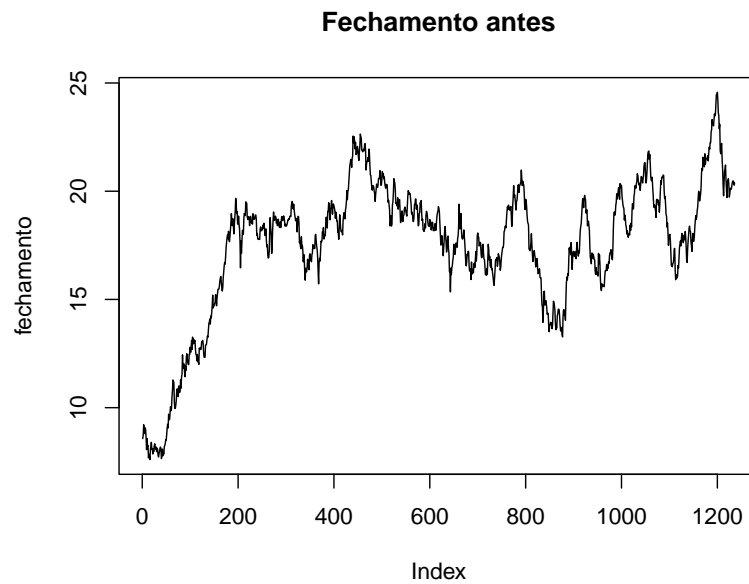
Fonte: O autor.

Ao aplicar a padronização podemos perceber que o conjunto de dados agora se encontra contido no intervalo 0 a 1 (Figuras 11 e 13 no eixo y e Figura 12 no eixo x).

1.5 Comparação

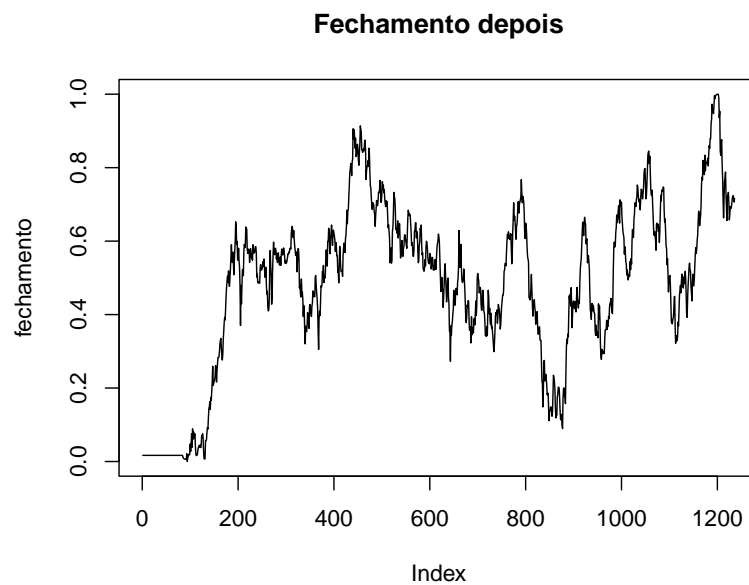
Nas imagens a seguir estão os gráficos onde os dados do fechamento estão presentes, na Figura 14 está os dados antes do tratamento e na Figura 15 estão os dados após sofrerem todos os processos que já foram mencionados neste trabalho.

Figura 14: Gráfico 14



Fonte: O autor.

Figura 15: Gráfico 15



Fonte: O autor.

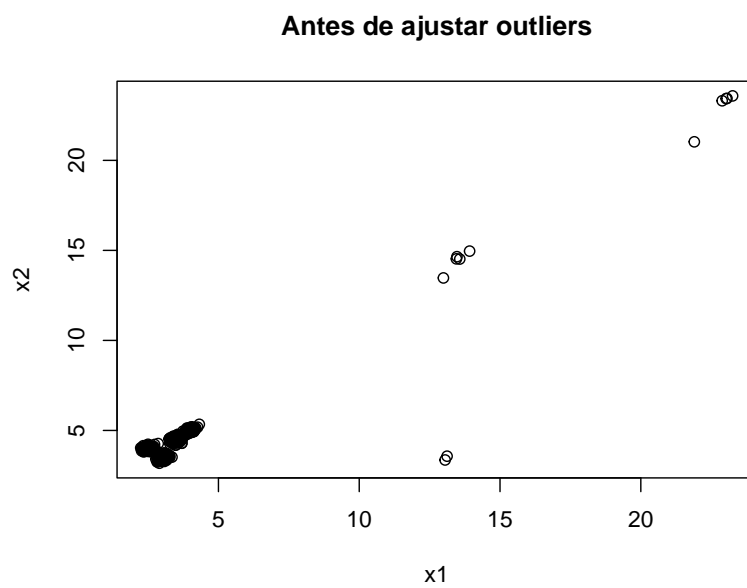
Note que além dos dados agora estarem contidos no intervalo 0 e 1 (eixo

y), também aparentam estarem um pouco mais próximos uns dos outros.

2 Questão 2

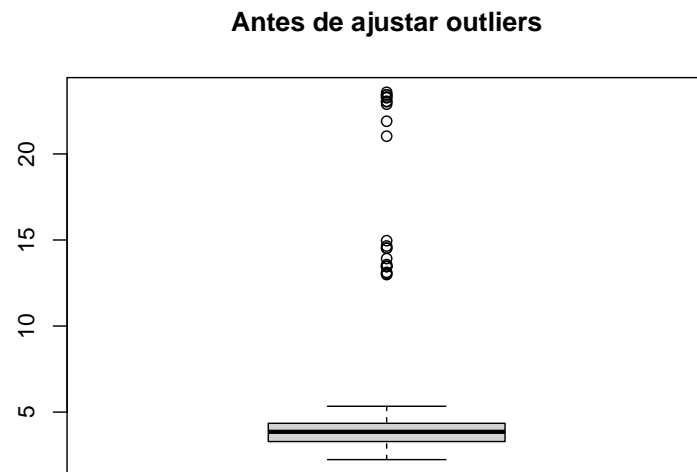
Os dados sem tratamento de amostr.csv nos gera os seguintes gráficos:

Figura 16: Gráfico 16



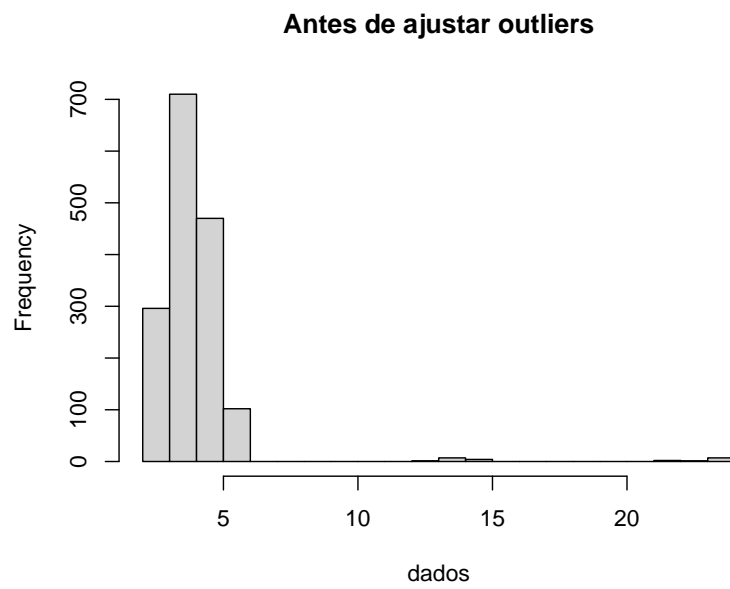
Fonte: O autor.

Figura 17: Gráfico 17



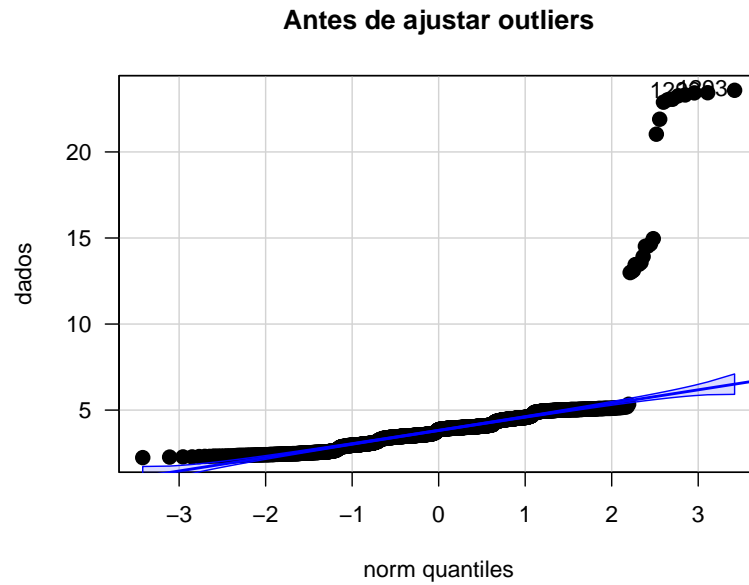
Fonte: O autor.

Figura 18: Gráfico 18



Fonte: O autor.

Figura 19: Gráfico 19



Fonte: O autor.

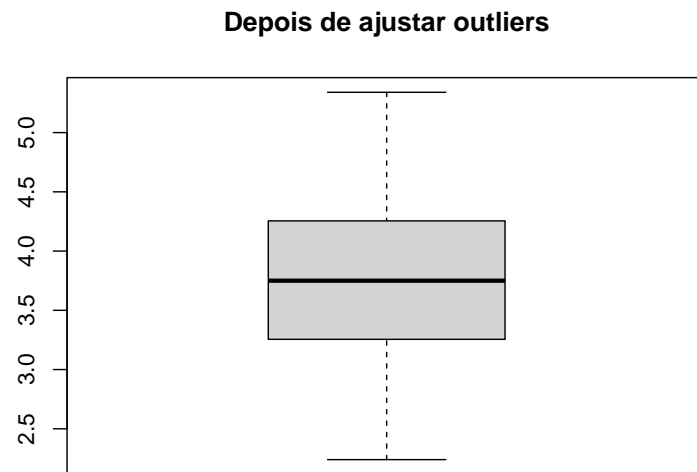
2.1 Tratamento de *outliers*

Para realizar o tratamento de *outliers* os seguintes passos serão seguidos:

- Visualização gráfica
- Detecção de *outliers*
- Normalização
- Padronização

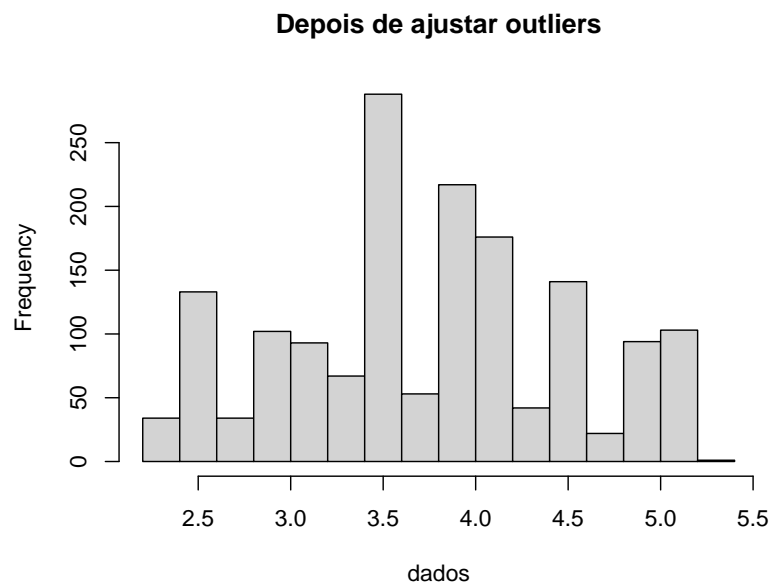
Desse modo, como no tópico 2.0 já visualizamos os gráficos e ao analisar-los é possível vermos *outliers*, vamos prosseguir para o ajuste de *outliers*:

Figura 20: Gráfico 20



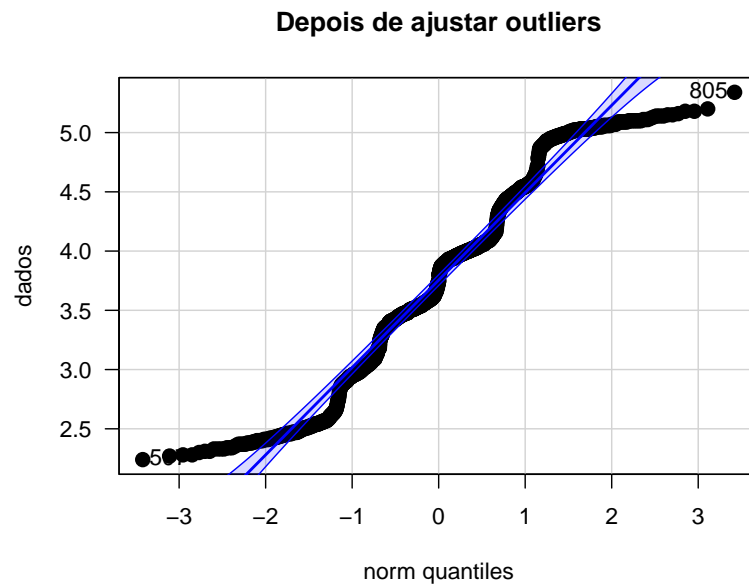
Fonte: O autor.

Figura 21: Gráfico 21



Fonte: O autor.

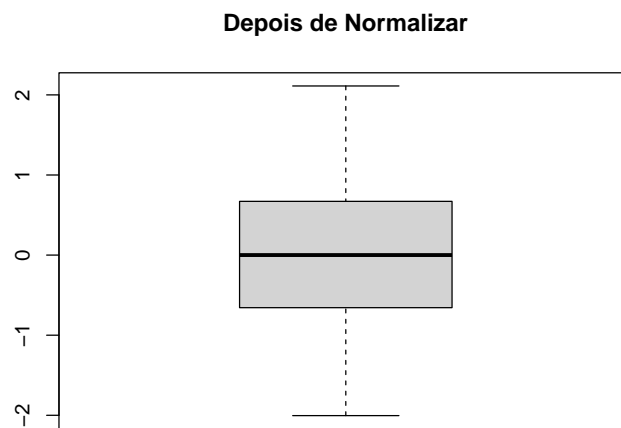
Figura 22: Gráfico 22



Fonte: O autor.

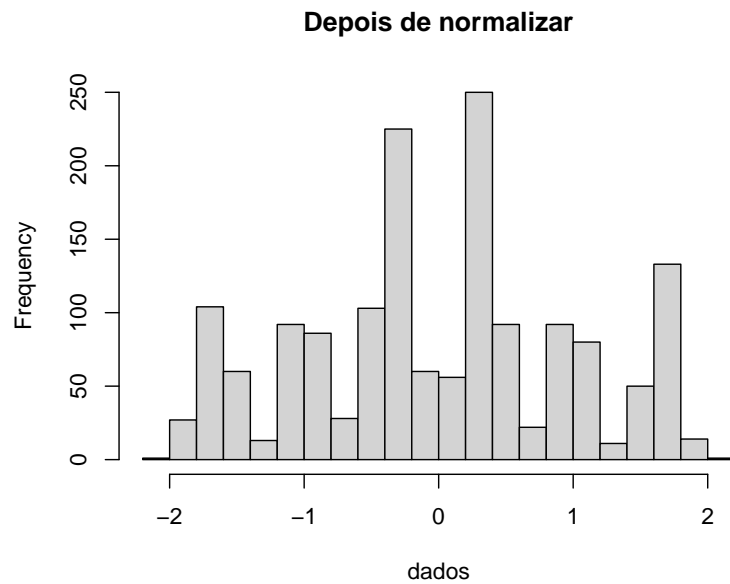
Em seguida realizaremos a Normalização dos dados e com isso produzirá os seguintes gráficos:

Figura 23: Gráfico 23



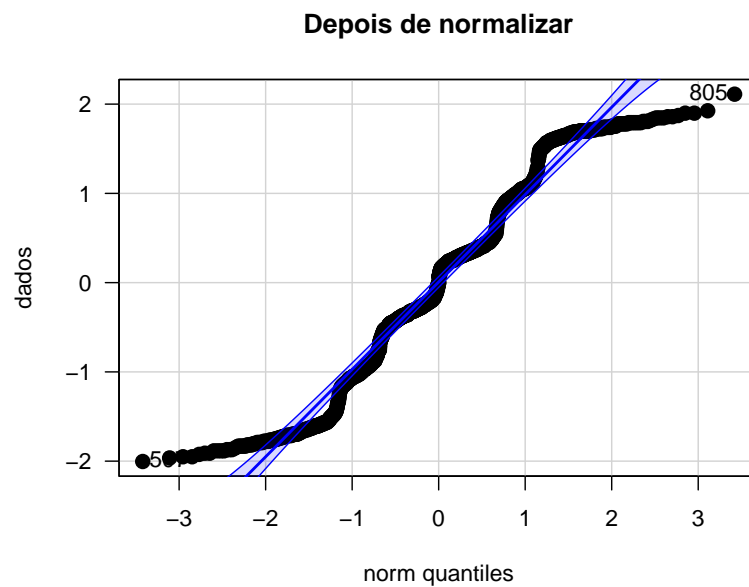
Fonte: O autor.

Figura 24: Gráfico 24



Fonte: O autor.

Figura 25: Gráfico 25

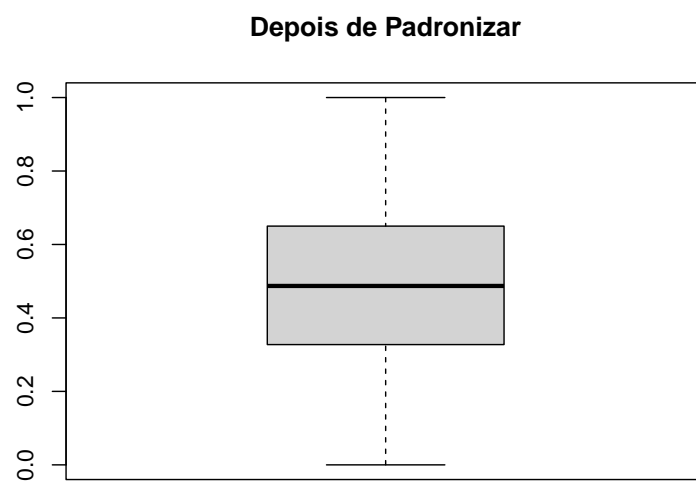


Fonte: O autor.

Por fim, aplicamos a padronização dos dados e teremos os seguintes grá-

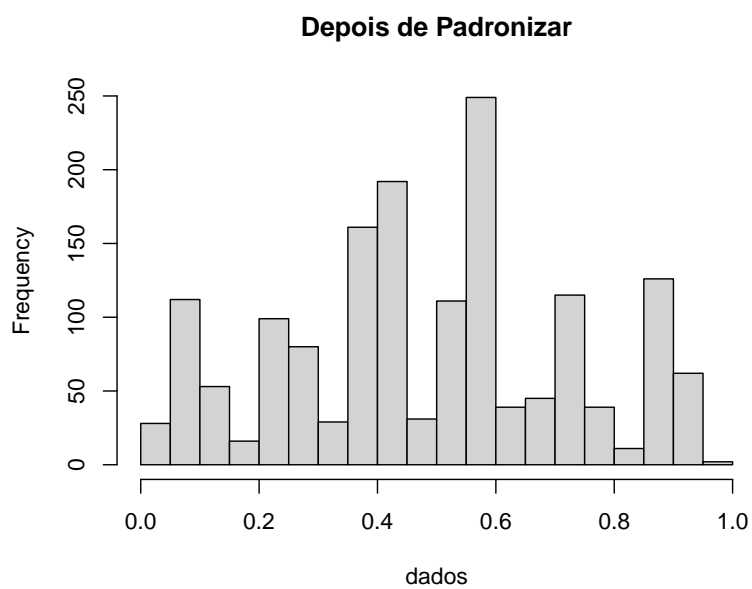
ficos:

Figura 26: Gráfico 26



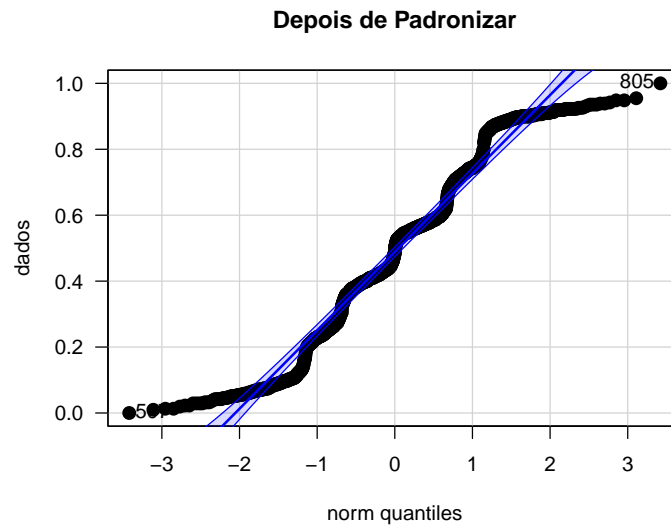
Fonte: O autor.

Figura 27: Gráfico 27



Fonte: O autor.

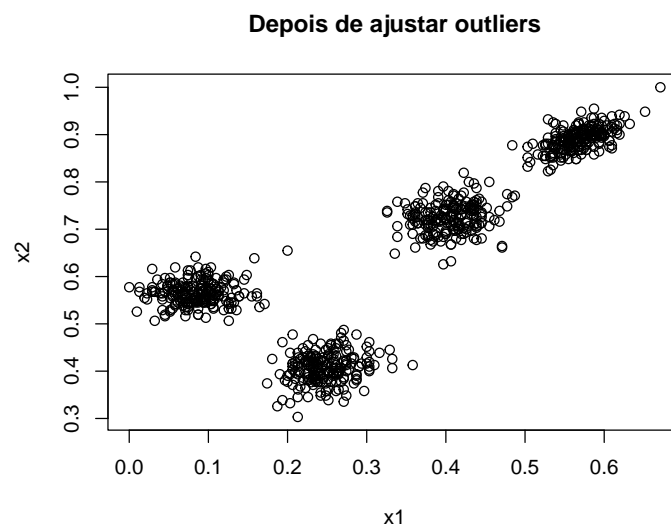
Figura 28: Gráfico 28



Fonte: O autor.

Portanto, após a realização de todas as etapas teremos o seguinte gráfico com os dados tratados:

Figura 29: Gráfico 29

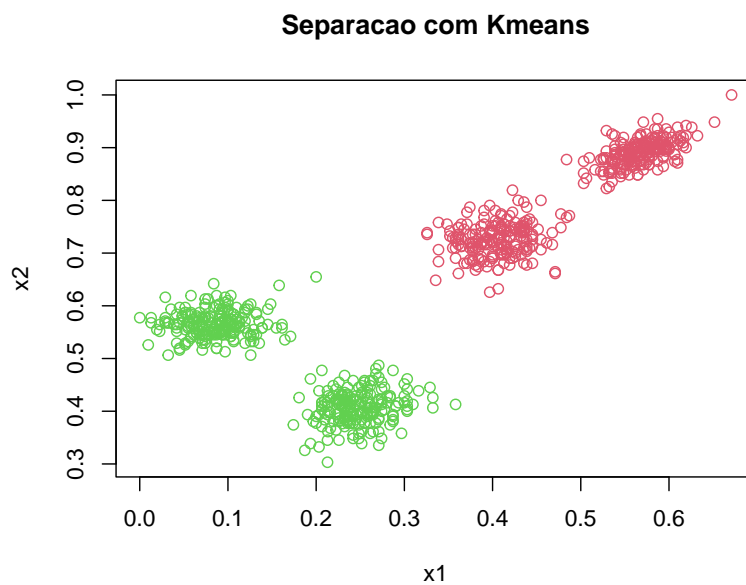


Fonte: O autor.

2.2 *K-means*

Aplicando a função *K-means* nos dados que foram tratados obtemos o seguinte gráfico:

Figura 30: Gráfico 30



Fonte: O autor.

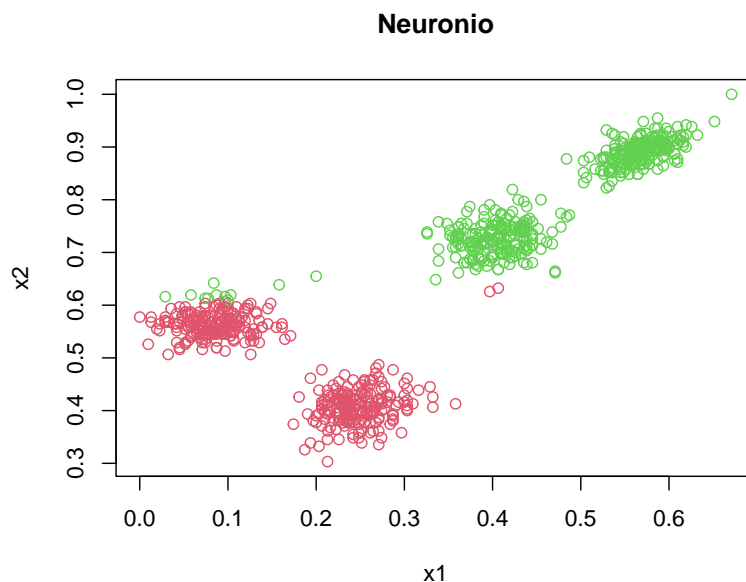
2.3 Neurônio

Para a realização desta etapa foram-se estipulados alguns pesos, são eles:

- $\text{Theta} = 1$
- $W1 = -0.12$
- $W2 = 1.5$
- $W3 = -0.9$

Com estes valores foi possível obter o seguinte gráfico (representa a melhor das tentativas feitas):

Figura 31: Gráfico 31



Fonte: O autor.

Em relação ao gráfico gerado utilizando a função *K-means* tem-se alguns pontos vermelhos e verdes que não ficaram em seus respectivos grupos. Porém, como a maior parte dos dados estão bem ajustados (em suas nuvens) considera-se aceitável este ajustes de peso manual.

3 Questão 3

O tratamento de dados antes do treinamento de uma Rede Neural Artificial (RNA) é uma etapa essencial no processo de construção de modelos de aprendizado de máquina. Envolve uma série de técnicas aplicadas aos dados brutos, com o objetivo de prepará-los para o treinamento da RNA.

Ao realizar este tratamento alguns pontos podem ser destacados:

- Melhorar a qualidade dos dados: Os dados brutos podem conter ruídos, inconsistências e valores atípicos que podem prejudicar o desempenho da RNA. O tratamento adequado dos dados ajuda a remover esses problemas, resultando em dados de melhor qualidade para o treinamento.
- Aumentar a eficiência do treinamento: Dados bem tratados podem acelerar o processo de treinamento da RNA.

- Evitar o viés nos dados: O viés ocorre quando certos padrões ou características são sobre ou sub-representados nos dados, o que pode levar a um modelo enviesado e resultados incorretos.

Em resumo, o tratamento de dados antes do treinamento de uma RNA é crucial para garantir a qualidade dos dados, acelerar o processo de treinamento, evitar o viés e preparar os dados para as necessidades específicas do modelo. Uma boa preparação dos dados contribui para a obtenção de resultados mais precisos e confiáveis por parte da RNA.