

# Metodos Quantitativos - A17

Marcelo Wzorek Filho, Rafael Leal Machado

June 2024

## 1 Introdução

Neste trabalho, realizarei uma análise estatística detalhada de duas variáveis: "normalized-losses" e "wheel-base", o objetivo principal é fornecer insights sobre essas variáveis, incluindo medidas de tendência central, dispersão e identificação de outliers. Para isso, seguirei um conjunto de etapas, que incluem:

1. Descrição estatística das variáveis, incluindo valor mínimo, valor máximo, média, desvio padrão, mediana, percentil 25 e percentil 75.
2. Identificação de outliers utilizando duas abordagens diferentes: distância interquartil (IQR) e limites baseados na média e desvio padrão.
3. Remoção dos outliers identificados e comparação dos gráficos de caixa antes e depois do recorte dos outliers.

O código fonte implementado para a resolução do exercício está disponível no [GitHub](#).

## 2 Normalized Losses

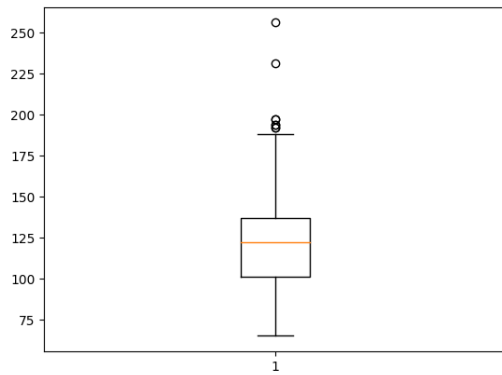
### 2.1 Descrição Estatística

- Valor mínimo: 65
- Valor máximo: 256
- Média: 113.83
- Desvio padrão: 37.35
- Mediana: 104
- Percentil 25: 93
- Percentil 75: 122

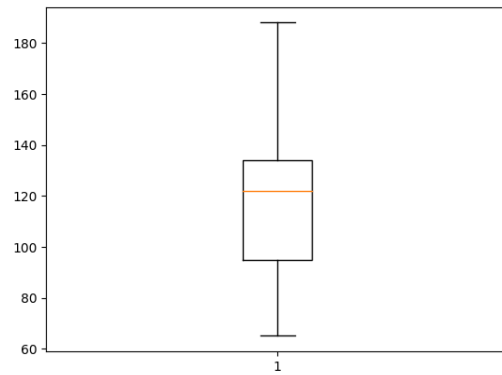
## 2.2 Outliers usando a Distância Interquartil (IQR)

- Limite inferior:  $93 - 1.5 \cdot (122 - 93) = 51.5$
- Limite superior:  $122 + 1.5 \cdot (122 - 93) = 163.5$
- Outliers: 164, 188, 192, 194, 197, 231, 256

## 2.3 Gráficos de Caixa



(a) Valores iniciais

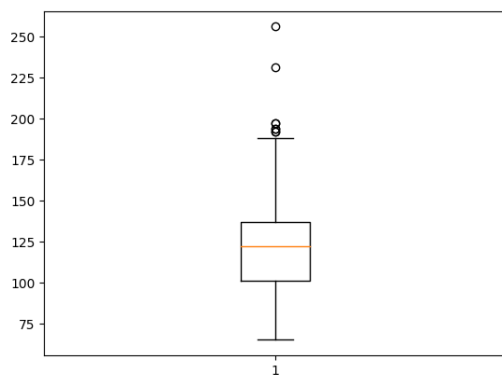


(b) Sem outliers (IQR)

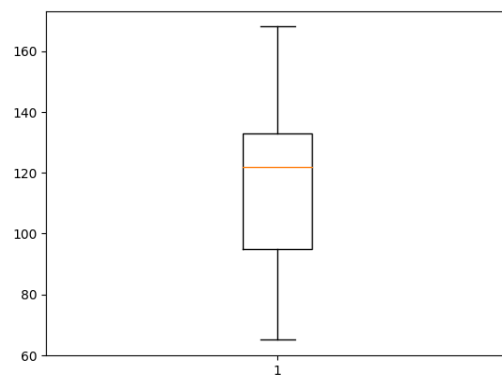
## 2.4 Outliers usando a Média e o Desvio Padrão

- Limite inferior:  $113.83 - 2 \cdot 37.35 = 39.13$
- Limite superior:  $113.83 + 2 \cdot 37.35 = 188.53$
- Outliers: 192, 194, 197, 231, 256

## 2.5 Gráficos de Caixa



(a) Valores iniciais



(b) Sem outliers (Média e DP)

### 3 Wheel Base

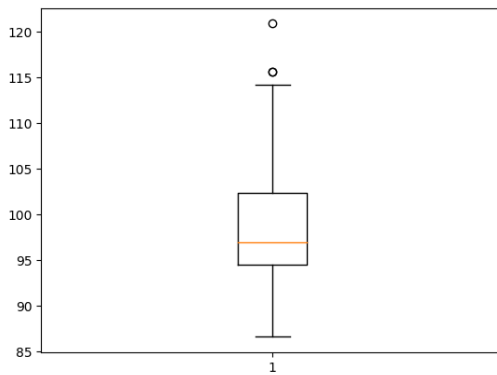
#### 3.1 Descrição Estatística

- Valor mínimo: 86.6
- Valor máximo: 120.9
- Média: 98.76
- Desvio padrão: 6.54
- Mediana: 96.3
- Percentil 25: 93.7
- Percentil 75: 102.4

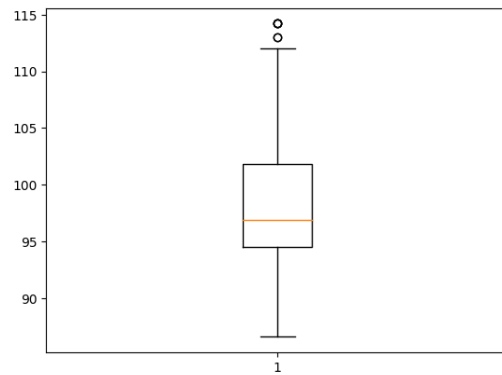
#### 3.2 Outliers usando a Distância Interquartil (IQR)

- Limite inferior:  $93.7 - 1.5 \cdot (102.4 - 93.7) = 80.05$
- Limite superior:  $102.4 + 1.5 \cdot (102.4 - 93.7) = 116.05$
- Outliers: 120.9

#### 3.3 Gráficos de Caixa



(a) Valores iniciais

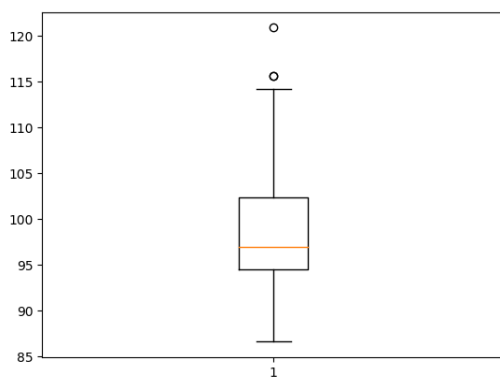


(b) Sem outliers (IQR)

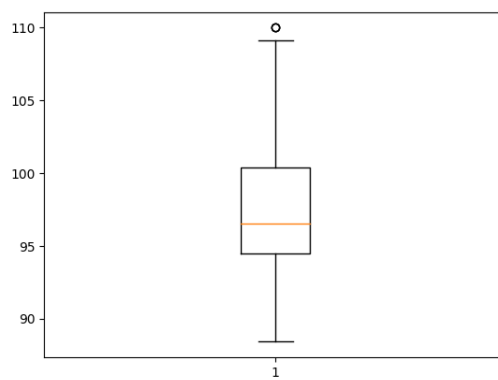
#### 3.4 Outliers usando a Média e o Desvio Padrão

- Limite inferior:  $98.76 - 2 \cdot 6.54 = 85.68$
- Limite superior:  $98.76 + 2 \cdot 6.54 = 111.84$
- Outliers: 120.9

### 3.5 Gráficos de Caixa



(a) Valores iniciais



(b) Sem outliers (Média e DP)