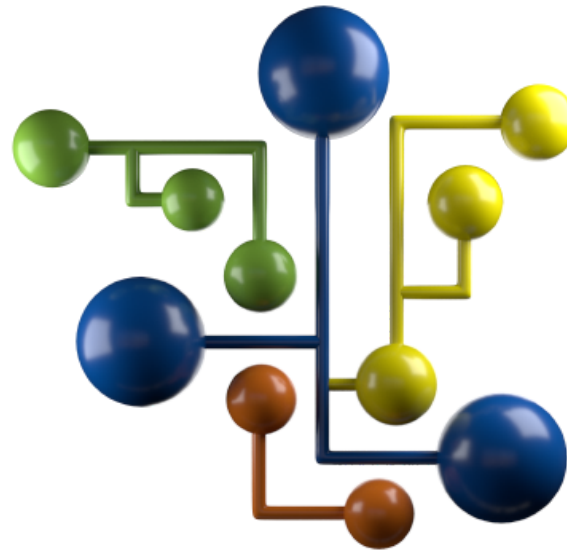


# Microsoft Power BI Para Business Intelligence e Data Science

## Lab 4

### Limpeza e Manipulação de Dados de Cadastro de Clientes no Power BI



Versão 1.0 - 02/04/23

Autor: Leonardo Simões

<https://www.datascienceacademy.com.br/course/microsoft-power-bi-para-business-intelligence-e-data-science>

# Índice

Capa

Índice

Descrição

Altura

Idade

Peso

Dashboard

# Descrição da Limpeza e Manipulação de Dados de Clientes

## Avaliação dos dados:

- . Tabela de contagem de ID de formas total e distinta;
- . Inspeção visual de valores;
- . Verificação dos valores para Altura;
- . Tabela de contagem para valores null de Idade e Peso.

## Problemas encontrados:

- . Valores duplicados de IDs;
- . Altura maior ou igual a 270;
- . Ausência de valores em Idade e Peso.

## Operações de Limpeza:

- . As linhas com IDs duplicados foram removidas, mantendo apenas uma das ocorrências;
- . As linhas com valores null em Idade e/ou Peso foram removidas;
- . As linhas com valores de Altura maior ou igual a 270 foram removidas.

fx = Table.SelectRows("#Linhas Agrupadas", each ([Contagem] <> 1))

|   | AB <sub>C</sub> ID_Cliente | 1 <sup>2</sup> <sub>3</sub> Contagem |
|---|----------------------------|--------------------------------------|
| 1 | 46                         | 2                                    |
| 2 | 55                         | 2                                    |
| 3 | 178                        | 3                                    |
| 4 | 185                        | 3                                    |

fx = Table.SelectRows("#Linhas Agrupadas", each ([Idade] = null))

|   | 1 <sup>2</sup> <sub>3</sub> Idade | 1 <sup>2</sup> <sub>3</sub> Contagem |
|---|-----------------------------------|--------------------------------------|
| 1 | null                              | 4                                    |

fx = Table.SelectRows("#Linhas Agrupadas", each ([Peso] = null))

|   | 1 <sup>2</sup> <sub>3</sub> Peso | 1 <sup>2</sup> <sub>3</sub> Contagem |
|---|----------------------------------|--------------------------------------|
| 1 | null                             | 4                                    |

fx = Table.Group("#Linhas Filtradas", {"Altura"}, {"Contagem", each Table.RowCount(\_), Int64.Type})

|   | 1 <sup>2</sup> <sub>3</sub> Altura | 1 <sup>2</sup> <sub>3</sub> Contagem |
|---|------------------------------------|--------------------------------------|
| 1 | 270                                | 1                                    |
| 2 | 278                                | 2                                    |

# Limpeza e Manipulação de Dados de Alturas de Clientes

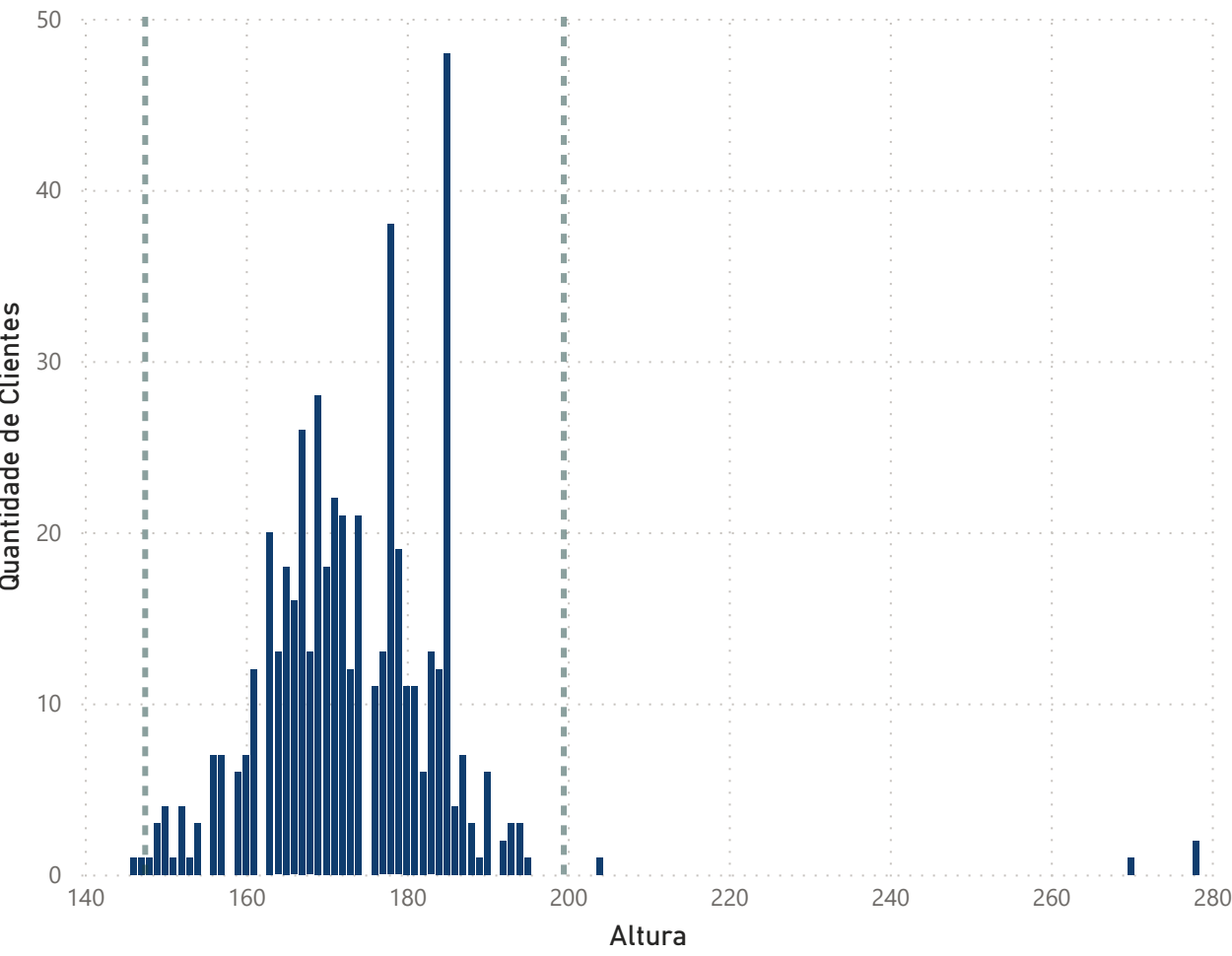
Estatísticas de Alturas do Dataset Original

| Mínimo | Limite Inferior | Q1     | Mediana | Média  | Q3     | Máximo | Limite Superior |
|--------|-----------------|--------|---------|--------|--------|--------|-----------------|
| 146    | 147,50          | 167,00 | 172,00  | 173,60 | 180,00 | 278    | 199,50          |

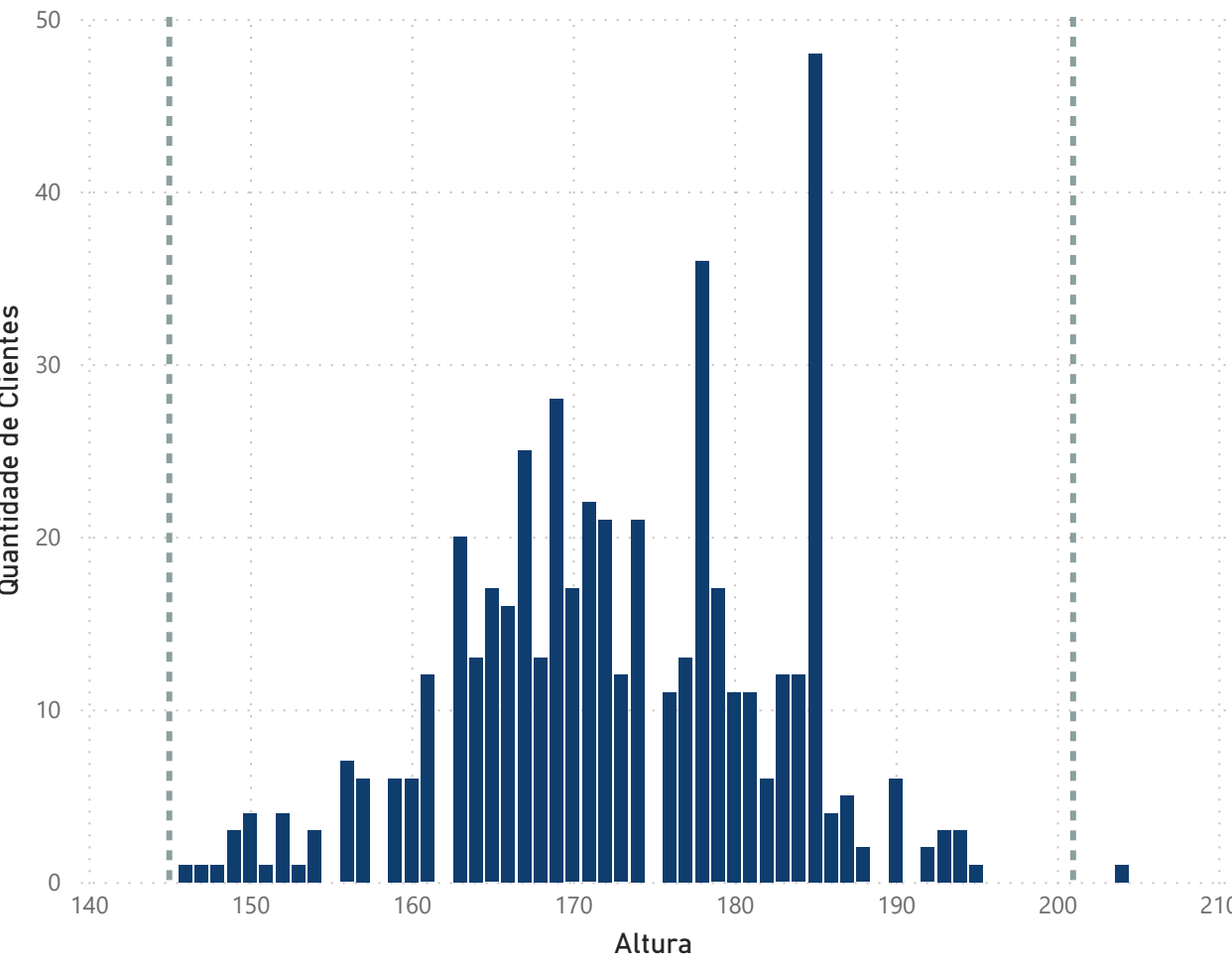
Estatísticas de Alturas do Dataset Corrigido

| Mínimo | Limite Inferior | Q1     | Mediana | Média  | Q3     | Máximo | Limite Superior |
|--------|-----------------|--------|---------|--------|--------|--------|-----------------|
| 146    | 145,00          | 166,00 | 172,00  | 172,89 | 180,00 | 204    | 201,00          |

Quantidade de Clientes por Altura (Original)



Quantidade de Clientes por Altura (Corrigido)

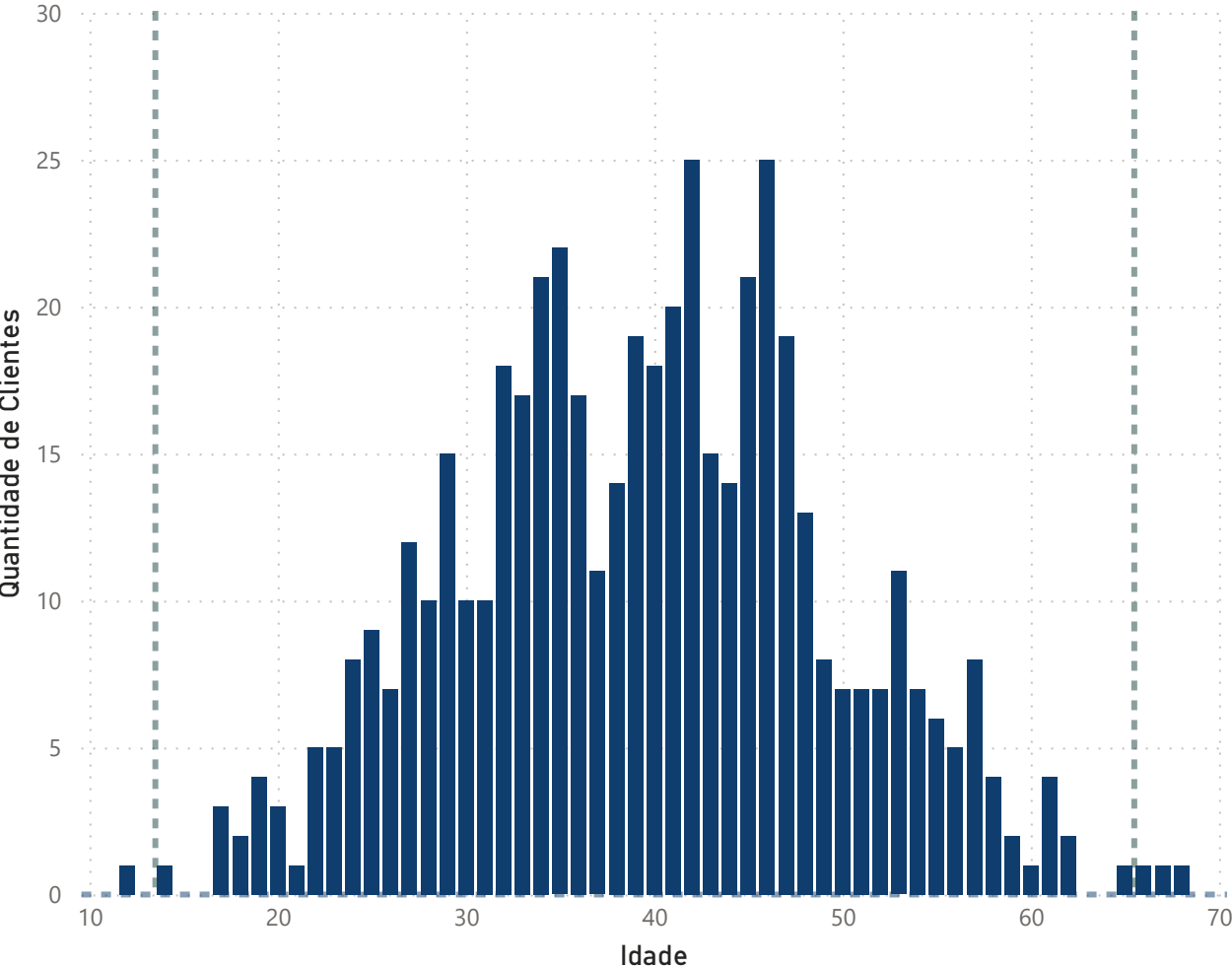


# Limpeza e Manipulação de Dados de Idades de Clientes

Estatísticas de Idade do Dataset Original

| Mínimo | Limite Inferior | Q1    | Média | Mediana | Q3    | Limite Superior | Máximo |
|--------|-----------------|-------|-------|---------|-------|-----------------|--------|
| 12     | 13,50           | 33,00 | 39,47 | 40,00   | 46,00 | 65,50           | 68     |

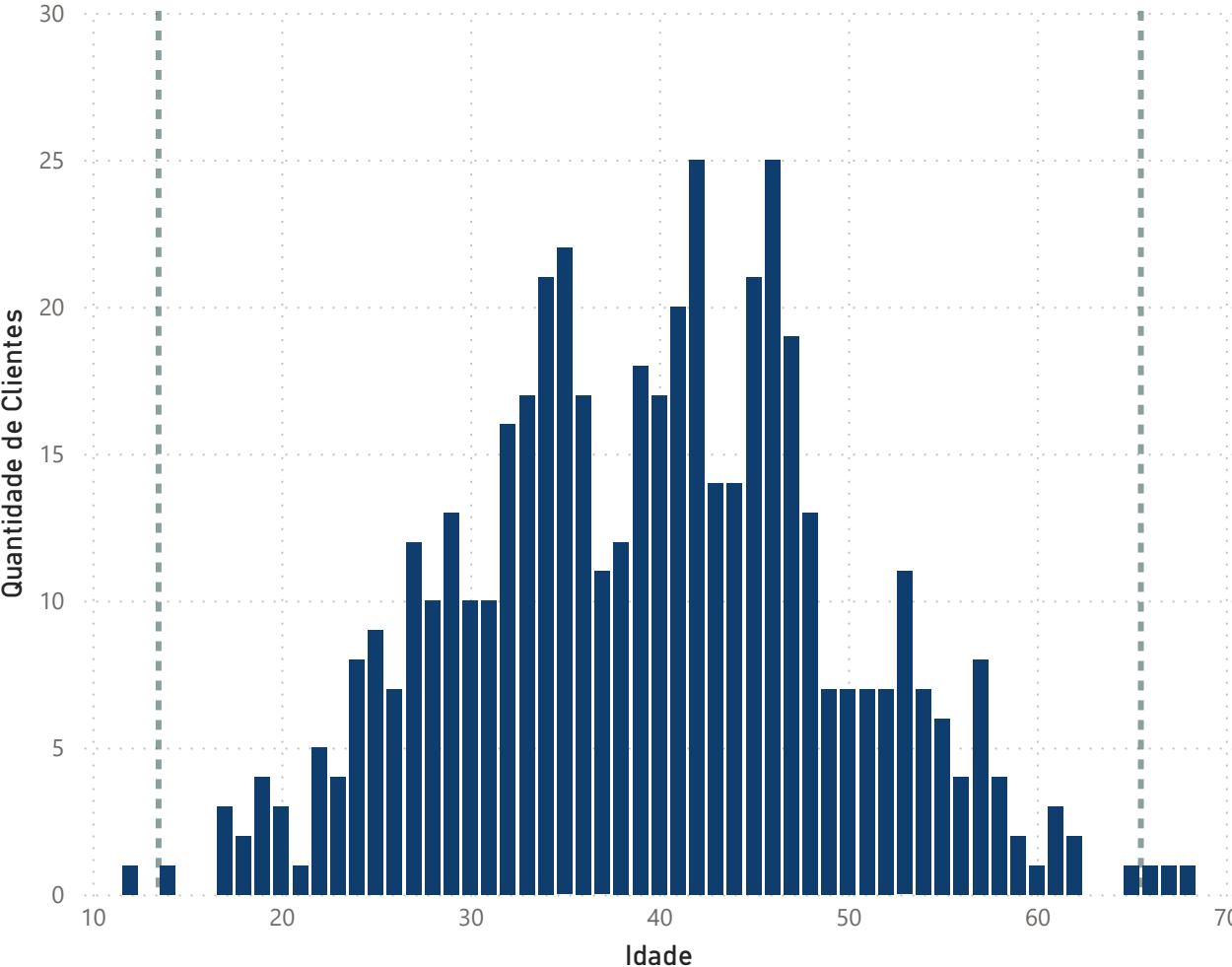
Quantidade de Clientes por Idade (Original)



Estatísticas de Idade do Dataset Corrigido

| Mínimo | Limite Inferior | Q1    | Média | Mediana | Q3    | Limite Superior | Máximo |
|--------|-----------------|-------|-------|---------|-------|-----------------|--------|
| 12     | 13,50           | 33,00 | 39,46 | 40,00   | 46,00 | 65,50           | 68     |

Quantidade de Clientes por Idade (Corrigido)

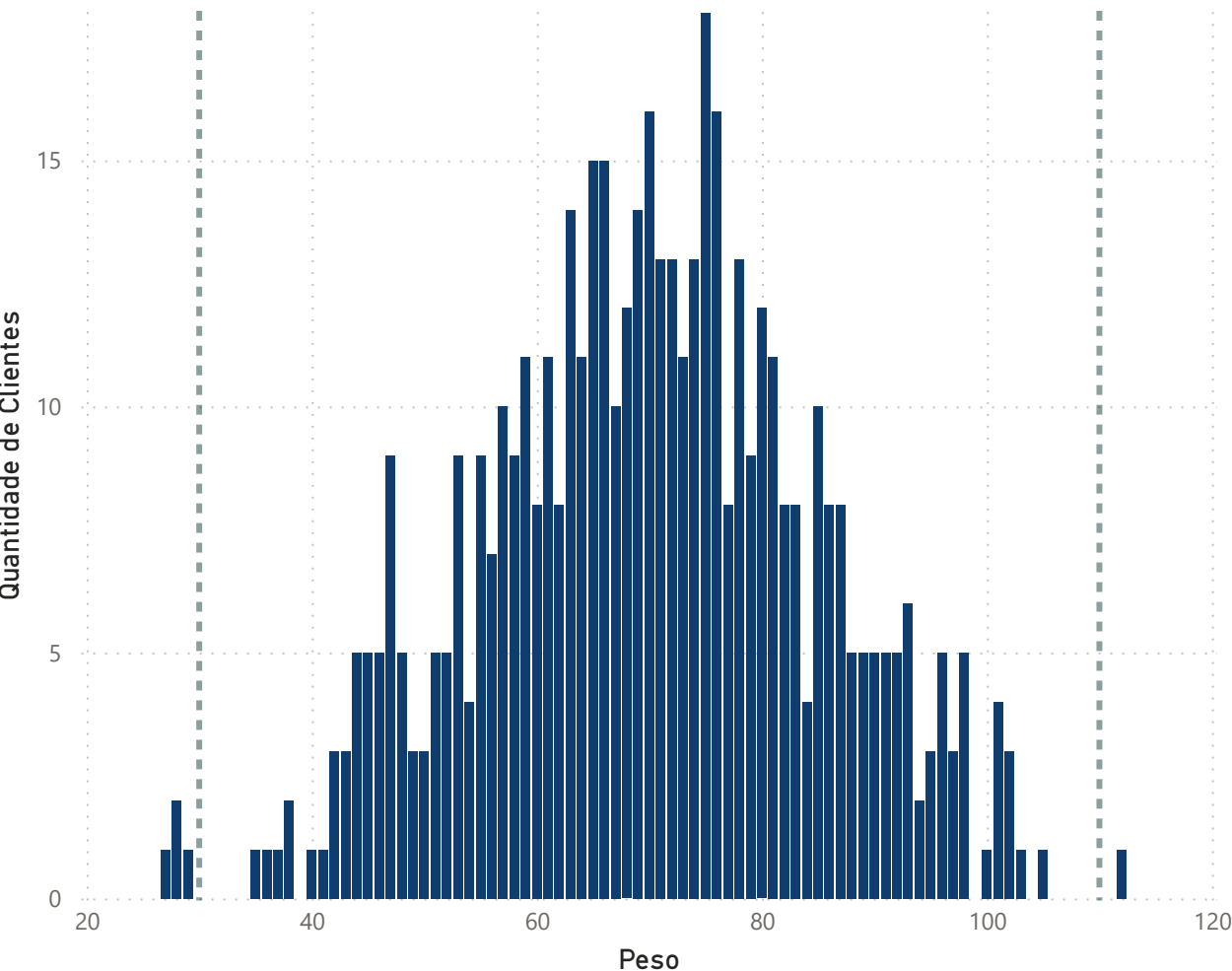


# Limpeza e Manipulação de Dados de Pesos de Clientes

Estadísticas de Pesos do Dataset Original

| Mínimo | Limite Inferior | Q1    | Média | Mediana | Q3    | Limite Superior | Máximo |
|--------|-----------------|-------|-------|---------|-------|-----------------|--------|
| 27     | 30,00           | 60,00 | 69,91 | 70,00   | 80,00 | 110,00          | 112    |

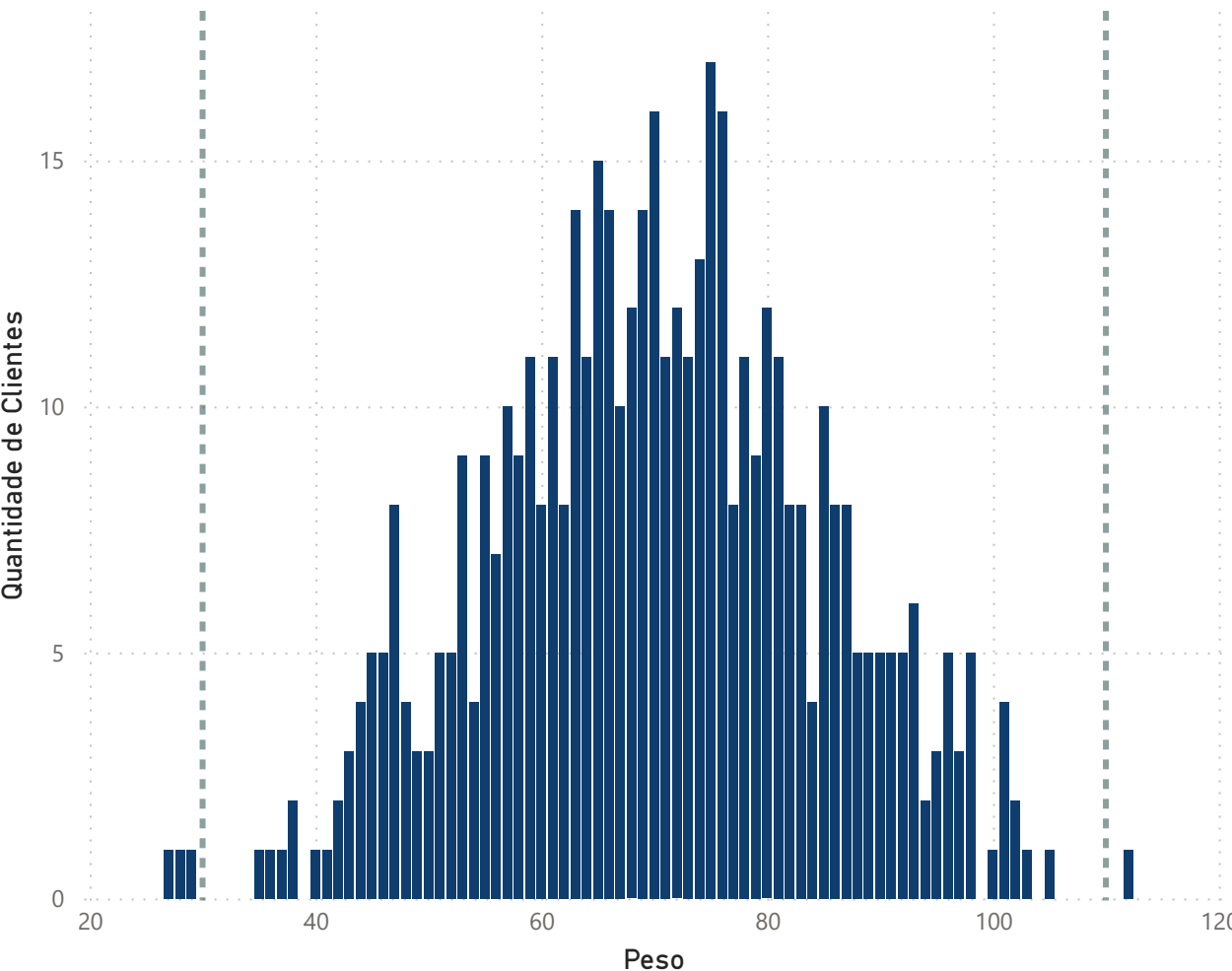
Quantidade de Clientes por Peso (Original)



Estadísticas de Pesos do Dataset Corrigido

| Mínimo | Limite Inferior | Q1    | Média | Mediana | Q3    | Limite Superior | Máximo |
|--------|-----------------|-------|-------|---------|-------|-----------------|--------|
| 27     | 30,00           | 60,00 | 70,09 | 70,00   | 80,00 | 110,00          | 112    |

Quantidade de Clientes por Peso (Corrigido)



# Dashboard de Clientes

70,09

Média de Peso

172,89

Média de Altura

39,47

Média de Idade

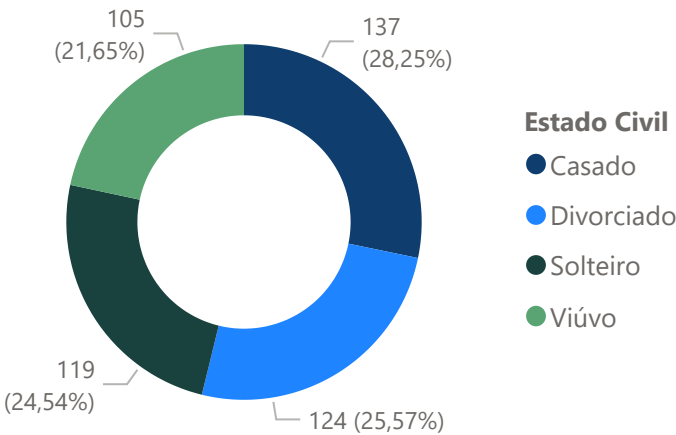
5,07 Mil

Média de Limite de Crédito

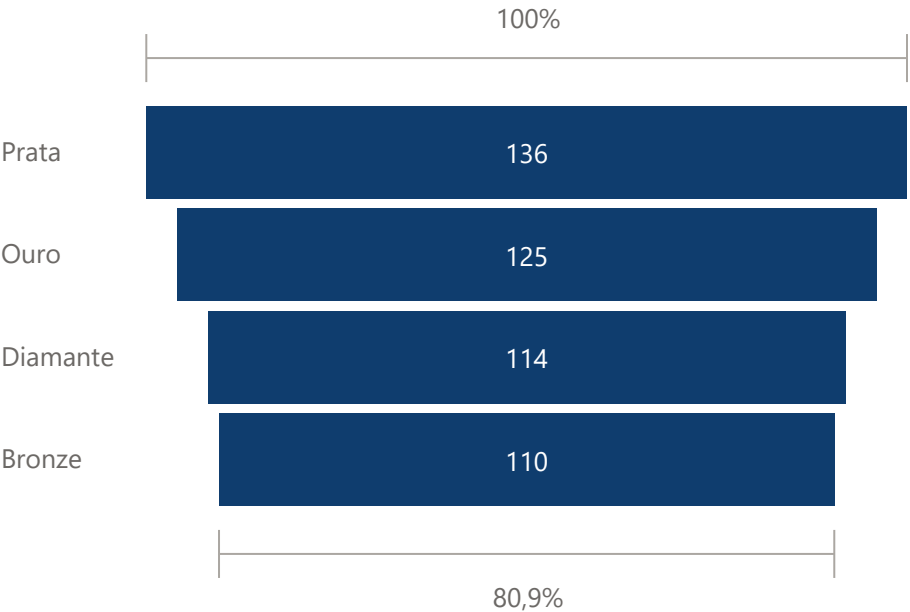
4,93

Média de Taxa de Desconto

Quantidade de Clientes por Estado Civil



Quantidade de Clientes por Tipo de Cliente



Média de Taxa de Desconto por Cidade

