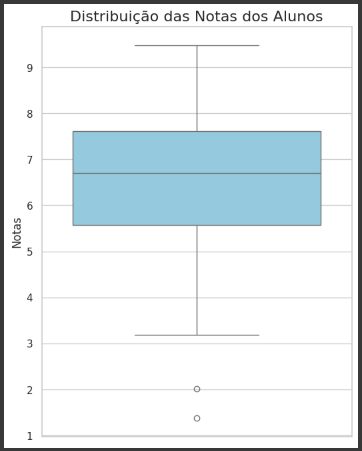
**Desafio 06**

1. **Crie o seu conjunto de dados rodando o seguinte código no google.colab:**

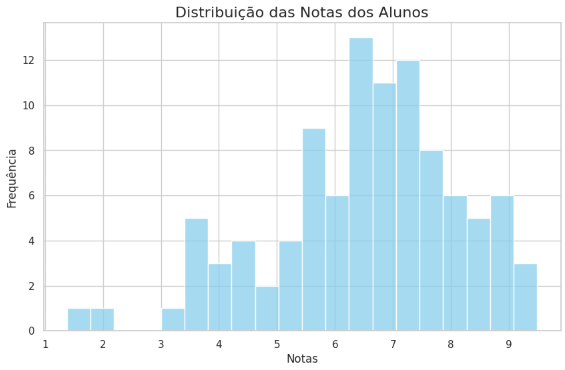
import numpy as np  
import pandas as pd  
  
# Gerando dados simulados  
np.random.seed(42)  
notas = np.concatenate([  
    np.random.normal(7, 1, 70),   # maioria dos alunos  
    np.random.normal(4, 1, 20),   # alguns com notas baixas  
    np.random.normal(9, 0.5, 10)  # e alguns com notas altas  
])  
  
# Garantir que todas as notas estejam entre 0 e 10  
notas = np.clip(notas, 0, 10)  
  
df = pd.DataFrame({'Nota': notas})  
df.to\_csv('notas\_alunos.csv', index=False)

**Obs: Esse código já irá salvar o dataset**notas\_alunos.csv.

1. **Plote um boxplot** para visualizar a dispersão e detectar possíveis outliers.



1. **Plote um histograma** da distribuição das notas.



1. **Encontre as medidas estatísticas** principais: média, mediana, mínimo e máximo.

**Resposta:**

* - Média: 6,5
* - Mediana: 6,69
* - Mínimo: 1,38
* - Máximo: 9,48

1. **Interprete os gráficos**:
   1. A distribuição é simétrica, assimétrica à esquerda ou à direita?

**Resposta:** Assimétrica à esquerda.

* 1. Há presença de outliers?

**Resposta:** Sim.

* 1. A média é maior ou menor que a mediana? O que isso indica?

**Resposta:** A média é menor que a mediana. Isso indica que há um grupo pequeno de alunos com notas baixas.

**Você deve entregar esse documento preenchido e o seu dataset notas\_alunos.csv.**