A **Correlação de Pearson** é uma medida que indica o grau de relação linear entre duas variáveis. O valor varia entre -1 e 1:

* **1**: Correlação positiva perfeita (ambas aumentam juntas).
* **0**: Nenhuma correlação linear.
* **-1**: Correlação negativa perfeita (uma aumenta enquanto a outra diminui).

**Como calcular a correlação de Pearson em Python?**

Você pode usar a biblioteca **NumPy** ou **Pandas** para isso, ou até a função específica do **scipy**.

Um valor de correlação de **0.7746** indica uma **correlação positiva forte** entre as duas variáveis.

### O que isso significa na prática?

* Conforme a variável **x** aumenta, a variável **y** tende a aumentar também.
* A relação é forte, mas não perfeita (que seria 1).
* Ainda pode haver alguma variação que não é explicada pela relação linear.

Base Teórica da Relação de Pearson

np.corrcoef(x, y)[0, 1]

O que ela faz?

np.corrcoef(x, y) — calcula a matriz de correlação de Pearson entre os vetores x e y.

Essa função retorna uma matriz 2x2:



​onde r é o coeficiente de correlação de Pearson entre x e y.

Por que [0, 1]?

[0, 1] acessa o elemento da primeira linha e segunda coluna da matriz, que é exatamente o valor do coeficiente de correlação entre x e y.

Testar com os seguintes valores:

x = np.array([-1, 0, 1])

y = np.array([1, 0, -1]) # y = -x (correlação -1)

# Agora vamos criar y que não tem correlação linear com x

x = np.array([-1, 0, 1])

y = np.array([1, 0, 1]) # y não é linearmente relacionada com x