

# Método Não supervisionado!

---

Deixa  
comigo!



# Não Supervisionada

- ▶ Na classificação não supervisionada, o analista não participa da definição dos parâmetros das classes (médias, variância, etc). E tarefa é deixada ao computador.
- ▶ Os algoritmos analisam o conjunto de dados disponíveis e nele identificam as classes mais frequentes. Cabe então ao usuário apenas determinar o número de classes que ele deseja obter.

1	2	1	4	12
2	1	2	3	15
1	3	3	5	10
1	1	3	6	12
1	2	3	9	16
21	21	20	21	26
22	22	24	23	27
23	29	29	30	29

Um grupo perto de 4  
e outro perto de 23 ?





# *Agrupamento (Clustering)*



# Agrupamento

---

- ▶ O método consiste em identificar nos dados grupos de pixels similares.
  - ▶ Cada grupo é descrito por seu vetor de médias. Estas médias são a princípio desconhecidas e devem ser achadas.
  - ▶ As médias devem ser tais que:
    - ▶ Os pixels de um mesmo grupos sejam similares entre si;
    - ▶ Pixels de grupos diferentes sejam diferentes;
  - ▶ Geralmente não se conhece *a priori* o número de clusters que existem na imagem.
  - ▶
-

# Similaridade

---

- ▶ Como media de similaridade pode ser usada a distância Euclidiana.
- ▶ Comprimento da reta que une o pixel e o centro de uma classe no espaço dos valores digitais.
- ▶ A Distância Euclidiana entre o pixel com leituras  $(I_0, J_0)$  e a média da classe  $(I_1, J_1)$  é:

$$D_e = \sqrt{(I_1 - I_0)^2 + (J_1 - J_0)^2}$$

---

# *k*-Means Clustering

---

- ▶ Como não se conhece os centros dos grupos, a busca da solução é iterativa.
  - ▶ Para começar, escolhe-se  $k$  centros (de maneira aleatória), um para cada grupo (clusters).
  - ▶ Com base na distância Euclidiana, verifica-se qual centro está mais próximo de cada pixel e assim se classificam os pixels. Isto associa cada pixel ao grupo mais próximo.
  - ▶ Como isto definiu os grupos, recalcula-se os verdadeiros centroides, pois já é conhecido quais pixels pertencem a cada grupo.
  - ▶ *Se houver diferença entre o centro (chute) inicial e o centro calculado, deve-se repetir o processo até atingir o centro verdadeiro.*
-

# Algoritmo *k-Means*

---

1. Determinar os centróides
  2. Atribuir a cada objeto do grupo o centróide mais próximo.
  3. Após atribuir um centróide a cada objeto, recalcular os centróides.
  4. Repetir os passos 2 e 3 até que os centróides não sejam modificados.
-

# Um exemplo em uma dimensão

---

Deixa  
comigo!



# Analizando os valores...

---

1	2	1	4	12
2	1	2	3	15
1	3	3	5	10
1	1	3	6	12
1	2	3	9	16
21	21	20	21	26
22	22	24	23	27
23	29	29	30	29

Quantos grupos?



# Algoritmo *k-Means*

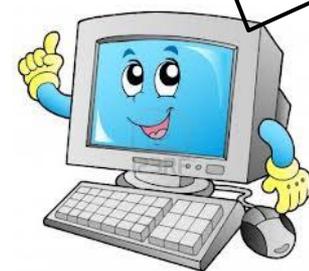
---

1. Determinar os centróides
  2. Atribuir a cada objeto do grupo o centróide mais próximo.
  3. Após atribuir um centróide a cada objeto, recalcular os centróides.
  4. Repetir os passos 2 e 3 até que os centróides não sejam modificados.
-

# Inicializar grupos atribuindo centros

1	2	1	4	12
2	1	2	3	15
1	3	3	5	10
1	1	3	6	12
1	2	3	9	16
21	21	20	21	26
22	22	24	23	27
23	29	29	30	29

*Não sei...  
Digamos um grupo  
perto 10 e outro  
perto de 20 ?*



Classif.

				2
2	2	2	2	2
2	2	2	2	2
2	2	2	2	2

medias = 4.2917 23.9375

Sinto muito, errou no chute!

Use as novas médias

# Repetir usando valores calculados

1	2	1	4	12
2	1	2	3	15
1	3	3	5	10
1	1	3	6	12
1	2	3	9	16
21	21	20	21	26
22	22	24	23	27
23	29	29	30	29

Classif.

				2
				2
2	2	2	2	2
2	2	2	2	2
2	2	2	2	2



medias = 3.8261 23.4118

Sinto muito, errou no chute!

Use as novas médias

# De novo...

1	2	1	4	12
2	1	2	3	15
1	3	3	5	10
1	1	3	6	12
1	2	3	9	16
21	21	20	21	26
22	22	24	23	27
23	29	29	30	29

Classif.

				2
				2
2	2	2	2	2
2	2	2	2	2
2	2	2	2	2

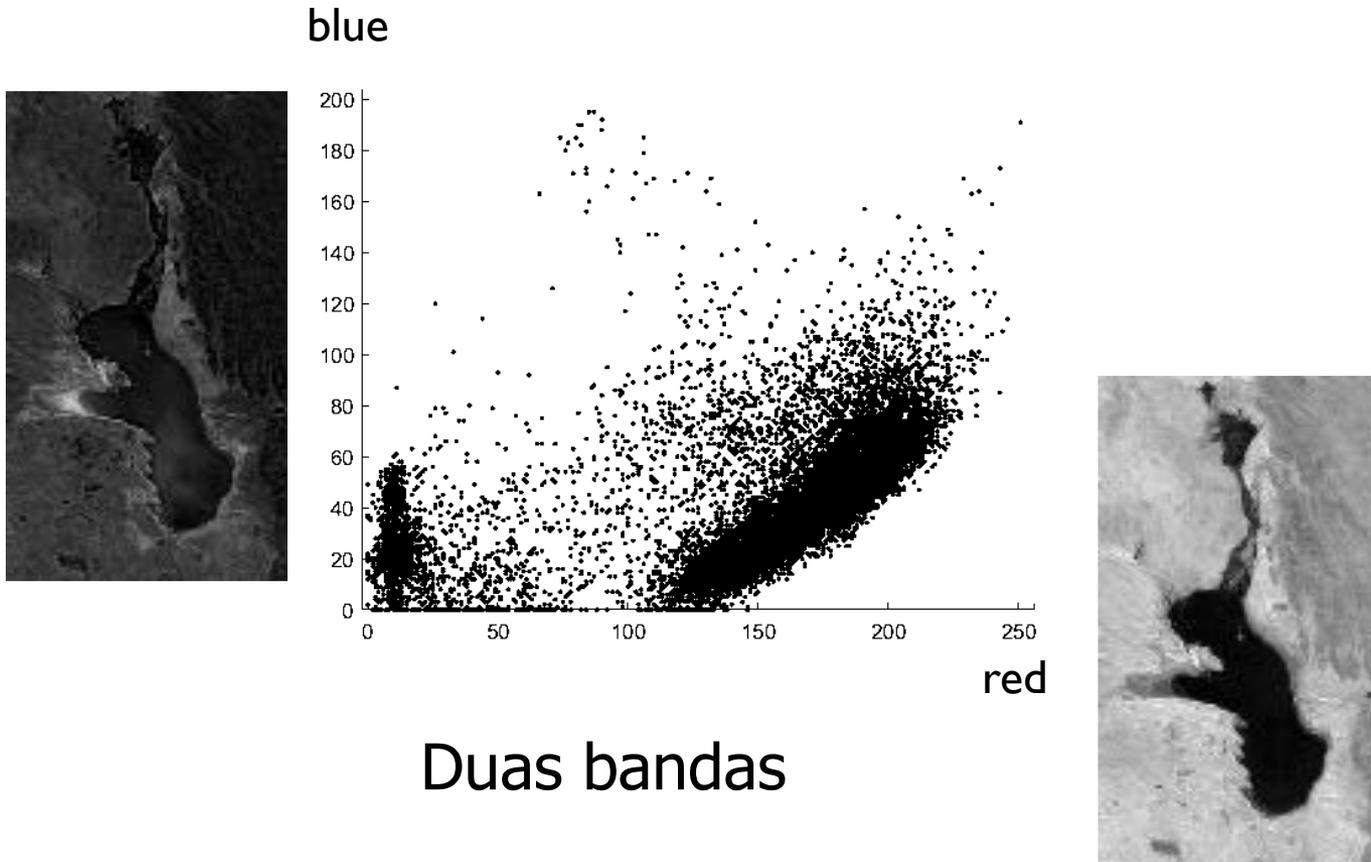


medias = 3.8261 23.4118

Acertou!

# *k*-Means – Exemplo

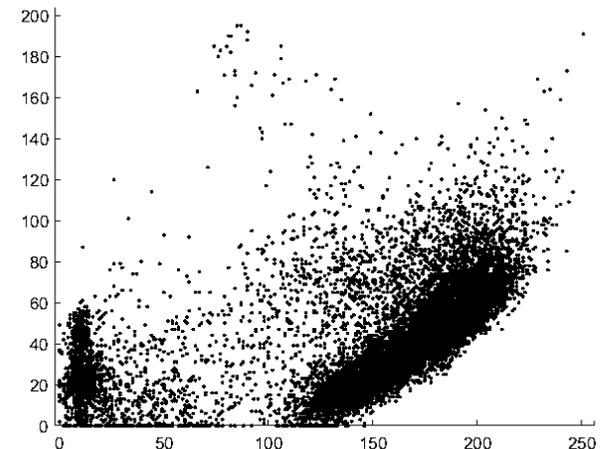
---



# definir

---

- ▶ A) quantos grupos você quer encontrar?
- ▶ Em nosso exemplo buscaremos 3 grupos.
  
- ▶ B) quando devemos parar? Quando a diferença entre os centroides atuais e os calculados for mesmo nula, ou quando atingir um pequeno valor?
  
- ▶ Digamos que quando a diferença for menor que 8.



# T=1

Inicie escolhendo centros  
arbitrários...

$C1 = [150 \ 50]$

$C2 = [100 \ 10]$

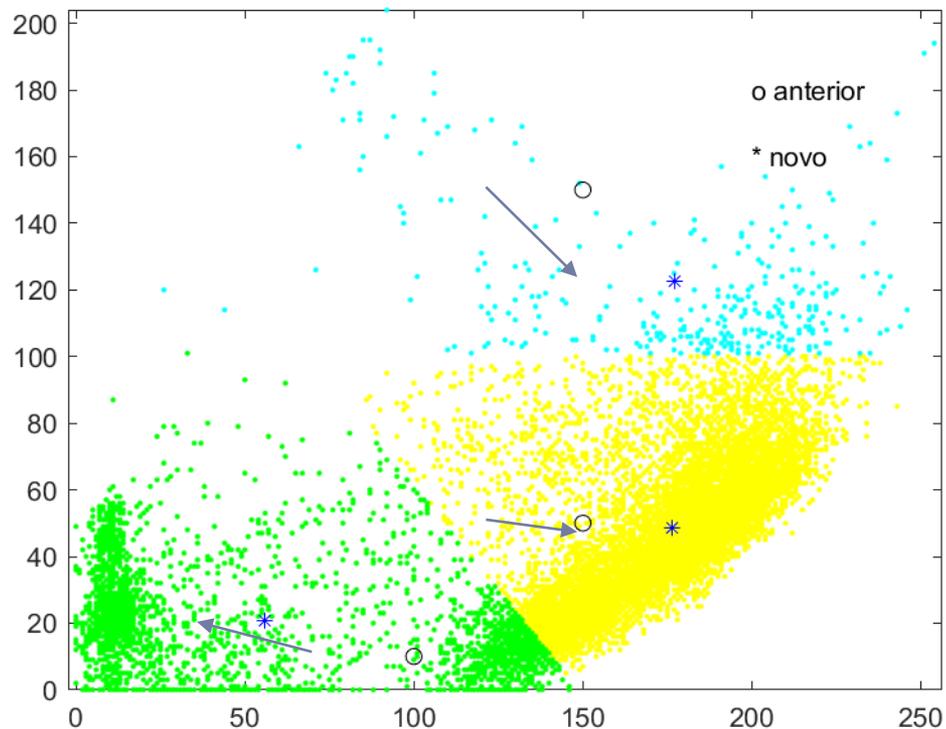
$C3 = [150 \ 150]$

Classifique a imagem

Calcule novo centro

Calcule a soma das diferenças  
entre os centros calculados e  
anteriores.

Iteração	distância
1	110.4



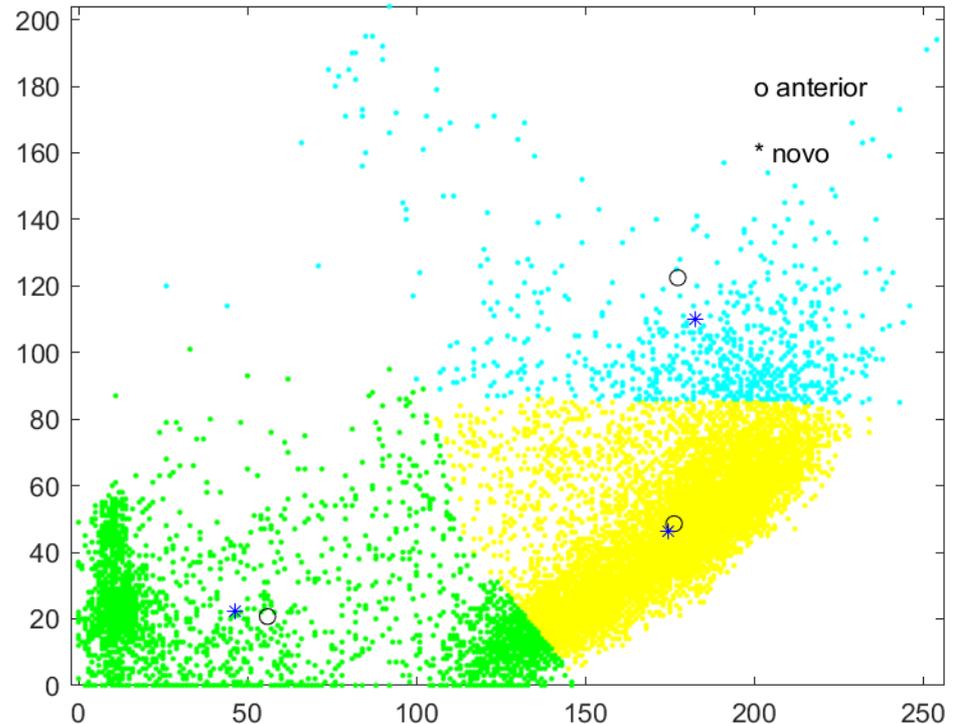
t=2

Se os centros dos grupos são diferentes dos anteriores, adote estes novos centros.

Note como os centros já são uma boa estimativa.

As diferenças entre os centros anteriores e os calculados são menores

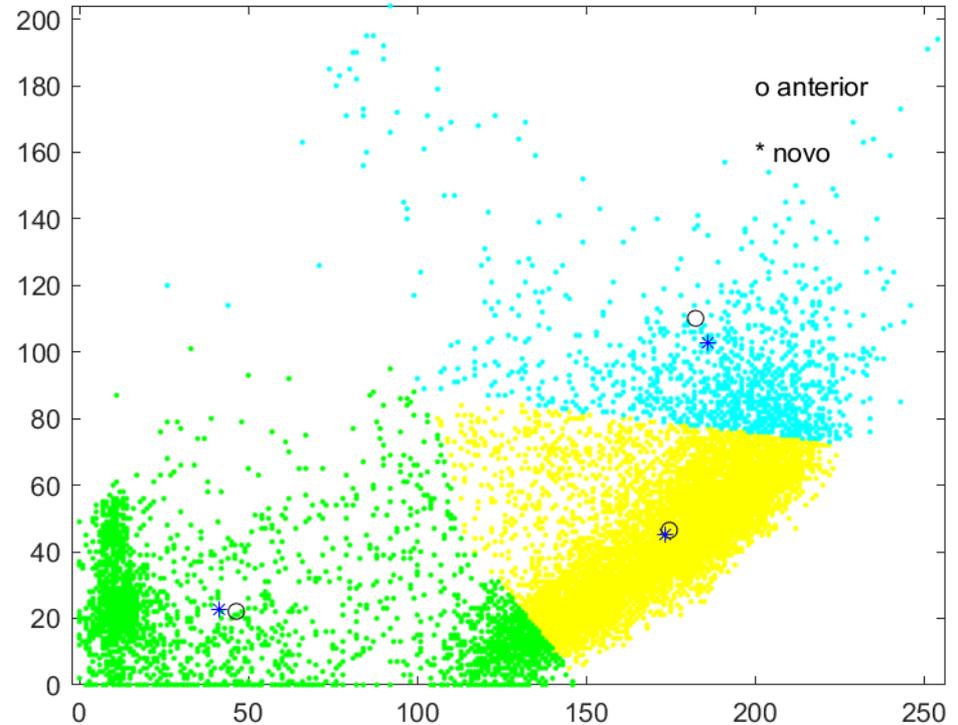
Iteração	distância
2	25.6



# T=3

E repita a  
classificação com  
estes centros

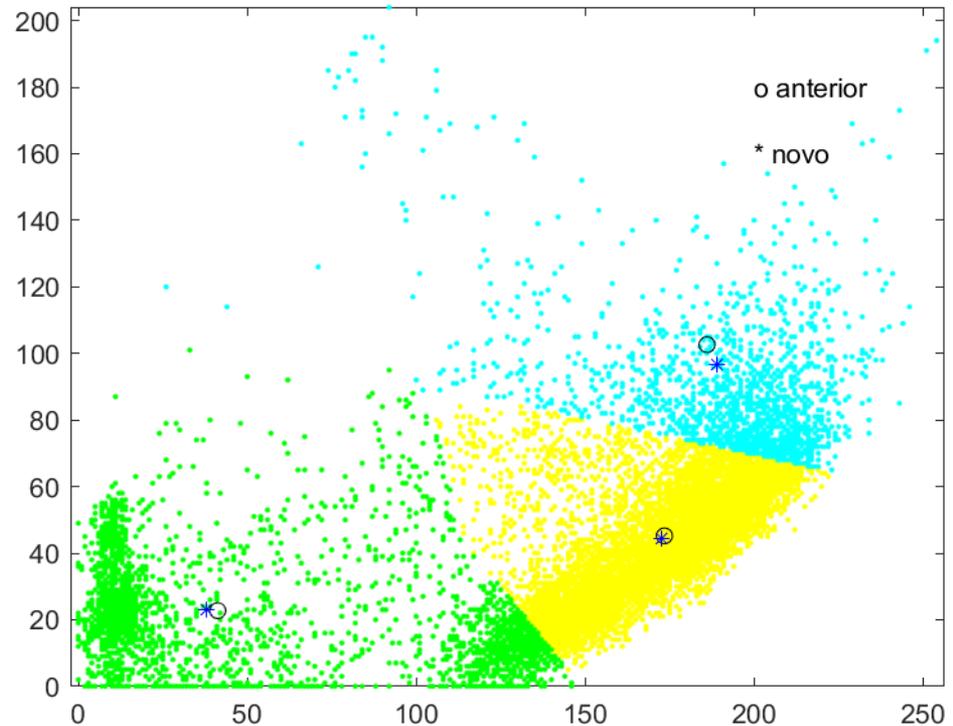
Cada vez são  
feitos ajustes  
menores



Iteração	distância
3	15.2

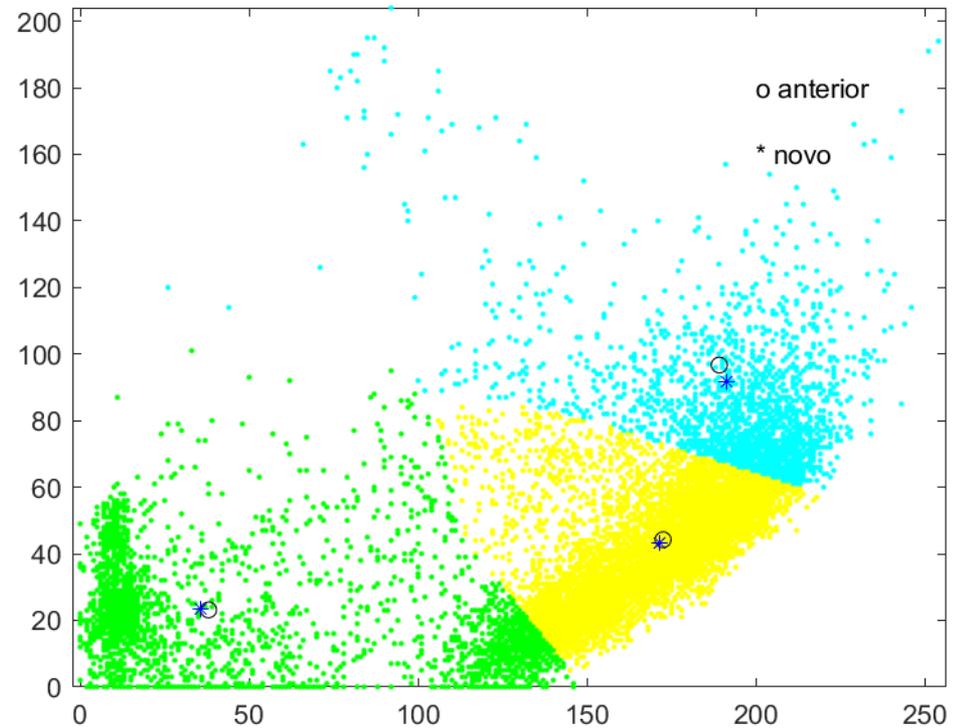
# T=4

Fazer isto até que os novos centros sejam próximos dos anteriores. Cada vez as correções são menores



Iteração	distância
4	11.3

# T=5

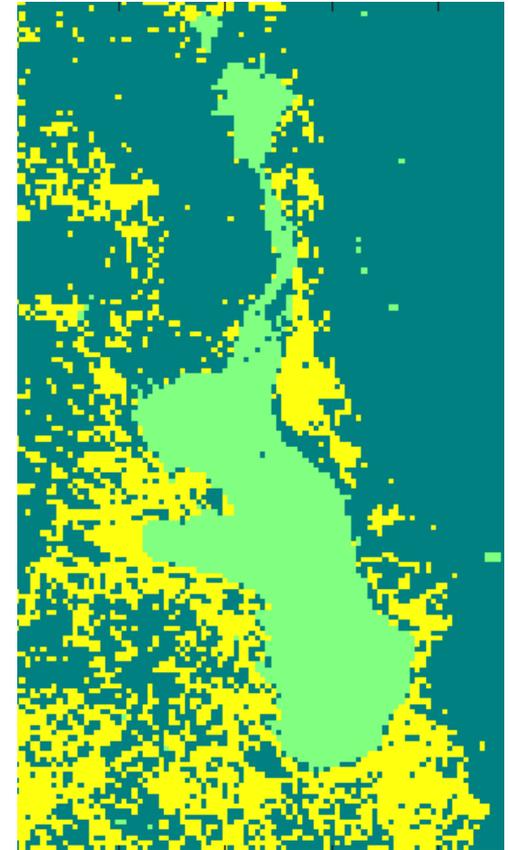
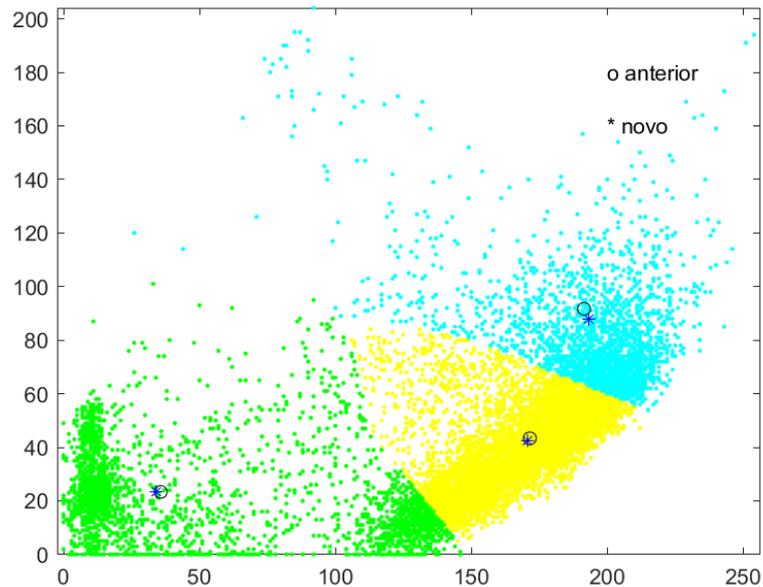


Iteração	distância
5	9.0
6	7.1

# Resumo

soma de distâncias entre centros usados e calculados

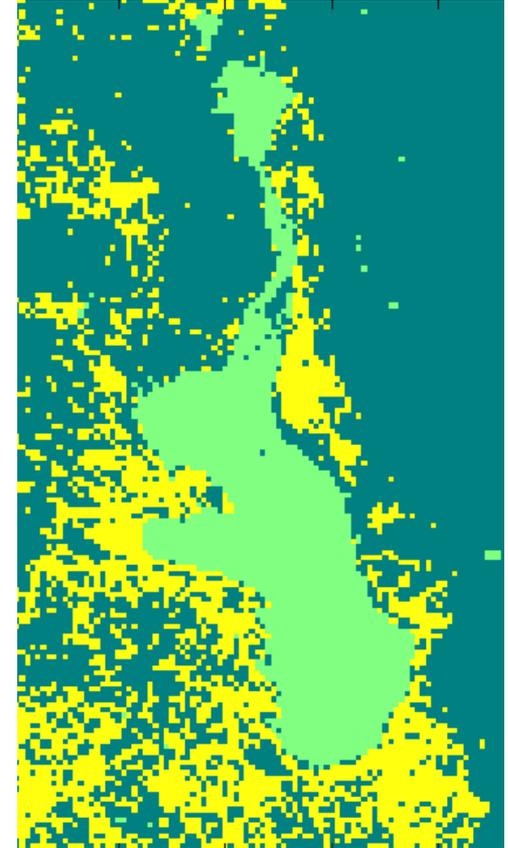
Iteração	distância
1	110.4
2	25.6
3	15.2
4	11.3
5	9.0
6	7.1



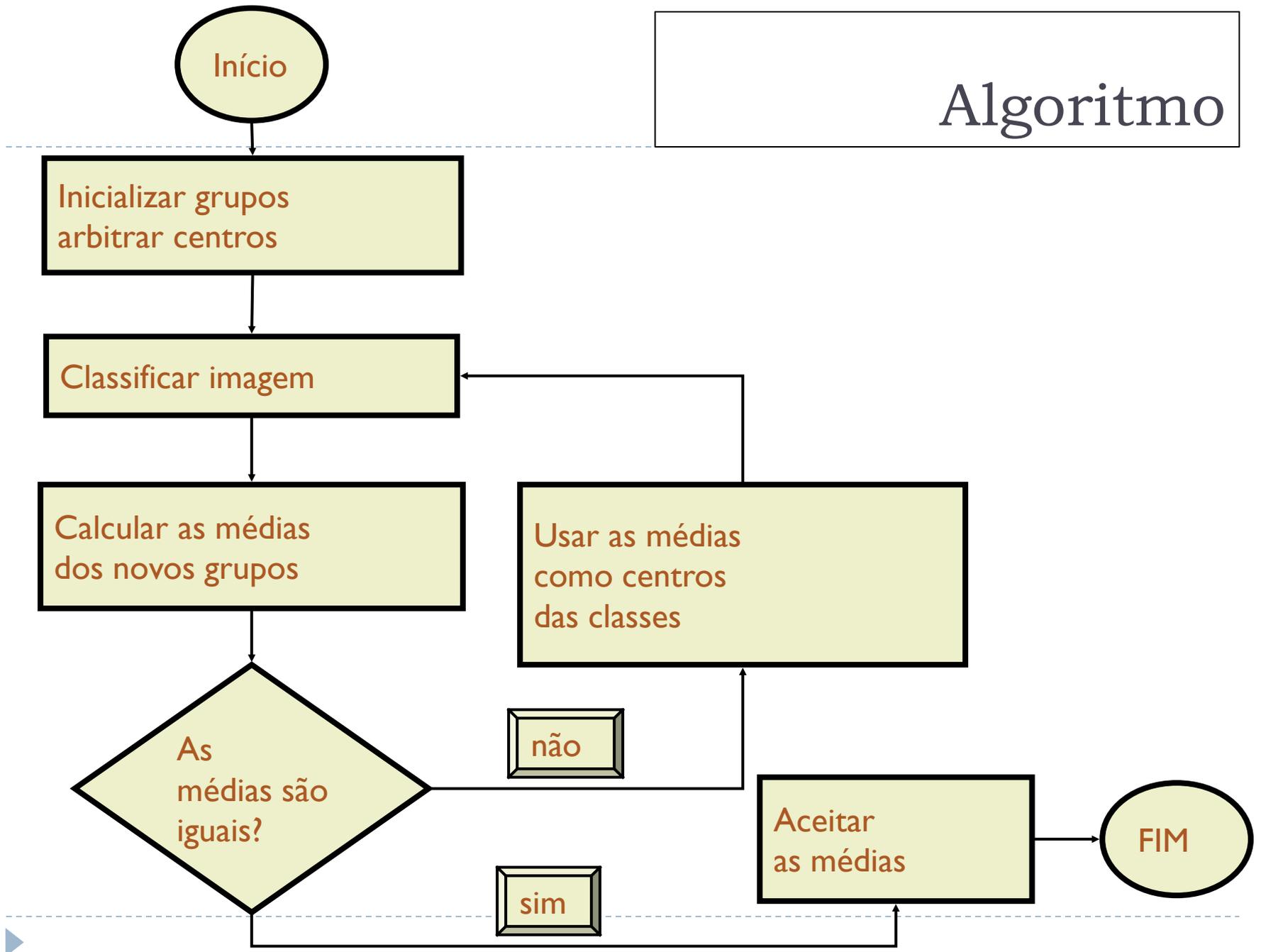
Resultado  
Classificação

# Para obter os segmentos...

- ▶ Após o agrupamento, deve se identificar os grupos de pixels conexos.
- ▶ Ex: Água
- ▶ Se for necessário, repetir o processo com mais grupos
- ▶ Grupos podem ser fundidos



# Algoritmo



# ISODATA

---

## **Iterative Self-Organizing Data Analysis Techniques.**

Este é uma variação do K-means que permite que o número de grupos seja ajustado automaticamente durante a iteração, combinando agrupamentos semelhantes e próximos, bem como dividindo aqueles com grandes valores de variância (dispersão).

Por isso, ele demanda a entrada de  
Número de grupos,  
Critério de convergência  
**Tamanho mínimo dos grupos**

---



# Não supervisionada

---

- ▶ Na classificação não supervisionada, os dados espectrais se superpõem à interpretação ou conhecimento prévio do usuário.
  - ▶ As classes não são modeladas de acordo à especificação do usuário, mas em vez disso usa-se abordagens estatísticas para dividir o espaço n-dimensional em grupos com a melhor separação possível.
  - ▶ Depois, pode-se atribuir significado a estes grupos, caso façam sentido no contexto da classificação
-