Parte 2

Parte 2.1 — Clustering com K-means (Manhattan, K=2)

1. Centróides Iniciais

- **C1** = P1 = (2, 3)
- **C2** = P4 = (8, 8)

2. Cálculo das Distâncias de Manhattan

Ponto	Distância até C1 (2,3)	Distância até C2 (8,8)	Cluster
P1	0	11	C1
P2	1	10	C1
P3	6	5	C2
P4	11	0	C2

3. Recalcular Centróides

Novo C1:

a. Média x = (2 + 3) / 2 = 2.5
b. Média y = (3 + 3) / 2 = 3.0
⇒ Novo C1 = (2.5, 3.0)

Novo C2:

c. Média x = (6 + 8) / 2 = 7.0
d. Média y = (5 + 8) / 2 = 6.5
⇒ Novo C2 = (7.0, 6.5)

4.

Cluster	Pontos atribuídos	Novo centróide
C1	P1 (2,3), P2 (3,3)	(2.5, 3.0)
C2	P3 (6,5), P4 (8,8)	(7.0, 6.5)

5. A escolha inicial influencia muito os agrupamentos. Maus centróides podem levar a agrupamentos errados ou necessidade de mais iterações.
6. Repetir: calcular distâncias dos pontos aos **novos centróides**, reatribuir clusters, recalcular centróides. Parar quando os clusters não mudam.
7. A forma de calcular a distância muda, o que pode alterar os agrupamentos, especialmente quando as diferenças de posição são diagonais.
8. Sim, os pontos (P1, P2) estão próximos, tal como (P3, P4). O agrupamento reflete bem a estrutura dos dados.
9. Deve-se usar um critério fixo, como escolher o primeiro centróide na ordem. Pode-se também recorrer ao critério da menor média histórica de distância ou randomizar.
10. Porque em cada iteração a função de custo (soma das distâncias aos centróides) **não aumenta**, e o número de agrupamentos possíveis é finito.

Parte 2.2 — Coeficiente de Silhueta (Manhattan)

1. e 2.

P1:

a(P1) = d (P1, P2) = |1-2| + |1-1| = 1

b(P1) = média{d (P1, P3), d (P1, P4)}

= (|1-4|+|1-4| + |1-5|+|1-4|)/2 = (6 + 7)/2 = **6.5**

s(P1) = (6.5 - 1)/6.5 ≈ **0.846**

P2:

a(P2) = d (P2, P1) = 1

b(P2) = média{d (P2, P3), d (P2, P4)}

= (|2-4|+|1-4| + |2-5|+|1-4|)/2 = (5 + 6)/2 = **5.5**

s(P2) = (5.5 - 1)/5.5 ≈ **0.818**

P3:

a(P3) = d (P3, P4) = |4-5| + |4-4| = 1

b(P3) = média{d (P3, P1), d (P3, P2)}

- = (|4-1|+|4-1| + |4-2|+|4-1|)/2 = (6 + 5)/2 = **5.5**

s(P3) = (5.5 - 1)/5.5 ≈ **0.818**

P4:

a(P4) = d (P4, P3) = 1

b(P4) = média{d (P4, P1), d (P4, P2)}

= (|5-1|+|4-1| + |5-2|+|4-1|)/2 = (7 + 6)/2 = **6.5**

s(P4) = (6.5 - 1)/6.5 ≈ **0.846**

3.

s médio = 0.846+0.818+0.818+0.846 / 4 = 3.328 / 4 ≈ 0.832