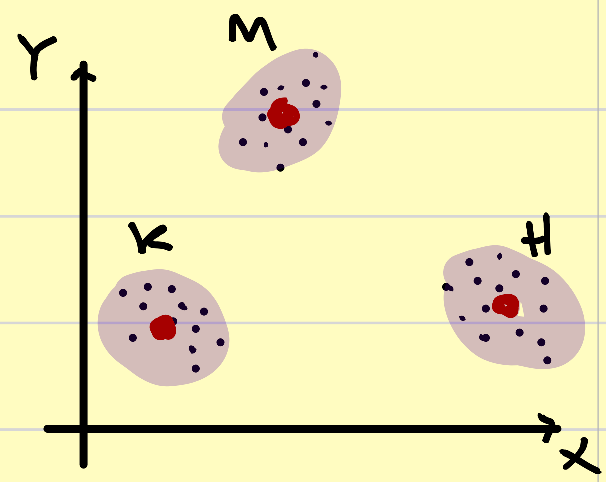
20251106_Statistik_WIN2

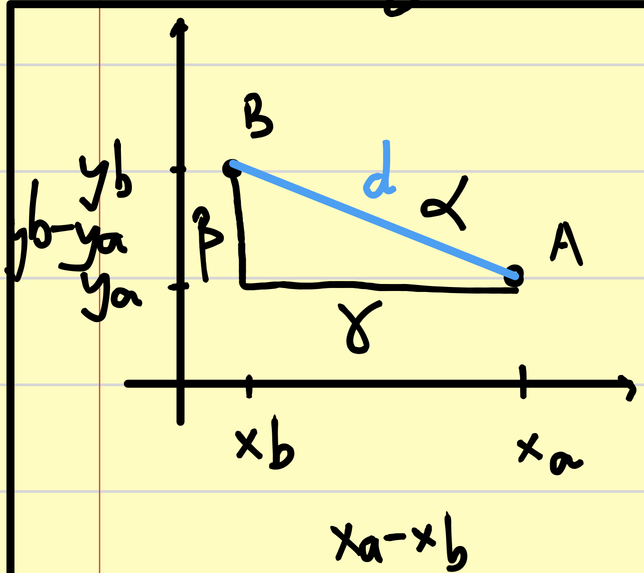
Maschinelles Lernen ...



K-means Clustering

Hypothese: Ähnliche Datensätze in der Nähe voneinander sind.

Dies setzt voraus, dass wir einen Abstand messen können.



$$d = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2}$$

$$\alpha^2 = \beta^2 + \gamma^2$$

"K" bedeutet die Anzahl Gruppen. (NACHTEIL: K muss vom Nutzer gegeben werden).

Schritte:

#0. Entscheidung über Anzahl Clusters. "K"
#1. Punkte werden in K Gruppen geteilt.
#2. Zentroide der Gruppen werden ermittelt.
#3. Abstand von den Zentroiden zu den Gruppen berechnen.
#4. Gruppenbildung nach dem geringsten Abstand.
#5. Zurück zum #1 bis die Gruppen konstant bleiben.

Beispiel. Gegeben sind die [X,Y] Koordinaten von 5 Werken. Bitte ermitteln Sie die optimale Position von 2 Läger, angenommen Alle Werke haben den gleichen Bedarf.

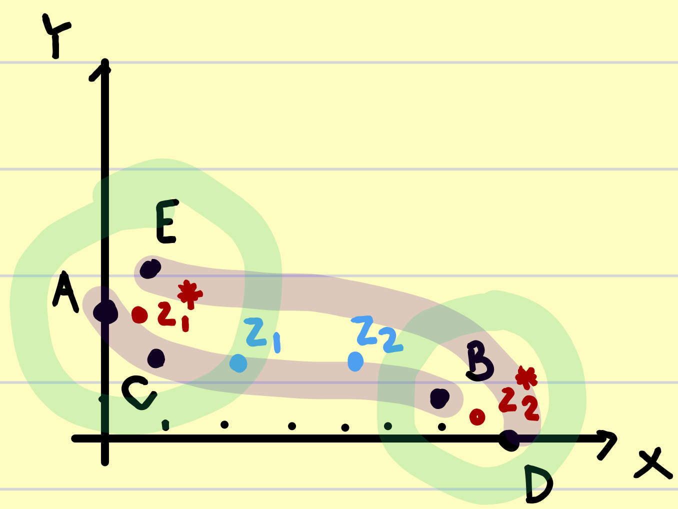
X: [0, 6, 1, 7, 2]

Y: [3, 1, 2, 0, 4]

A B C D E

Anfangsgruppen: Gruppe 1. [A,B,C] 2. [D,E].

#2. $z_1\left[\frac{0+6+1}{3}, \frac{3+1+2}{3}\right] = [2'33, 2]$ $z_2\left[\frac{7+2}{2}, \frac{0+4}{2}\right] = [4'5, 2]$



#3. $d_{A,z_1} = \sqrt{(0-2'33)^2+(3-2)^2} = 2'53$; $d_{A,z_2} = \sqrt{(0-4'5)^2+(3-2)^2} = 4'6$

$d_{B,z_1} = \sqrt{(6-2'33)^2+(1-2)^2} = 3'8$; $d_{B,z_2} = \sqrt{(6-4'5)^2+(1-2)^2} = 1'8$

$d_{C,z_1} = \sqrt{(1-2'33)^2+(2-2)^2} = 1'33$; $d_{C,z_2} = \sqrt{(1-4'5)^2+(2-2)^2} = 3'5$

$d_{D,z_1} = \sqrt{(7-2'33)^2+(0-2)^2} = 5'08$; $d_{D,z_2} = \sqrt{(7-4'5)^2+(0-2)^2} = 3'2$

$d_{E,z_1} = \sqrt{(2-2'33)^2+(4-2)^2} = 2'027$; $d_{E,z_2} = \sqrt{(2-4'5)^2+(4-2)^2} = 3'2$

Gruppen #1*: [A,C,E] #2*: [B,E]

#2. $z_1^*\left[\frac{0+1+2}{3}, \frac{3+2+4}{3}\right] = [1,3]$ $\quad z_2^*\left[\frac{6+7}{2}, \frac{1+0}{2}\right] = [6'5, 0'5]$

#3.

$d_{A,z_1^*} = \sqrt{(0-1)^2+(3-3)^2} = 1$; $d_{A,z_2^*} = \sqrt{(0-6'5)^2+(3-0'5)^2} = 6'9$

$d_{C,z_1^*} = \sqrt{(1-1)^2+(2-3)^2} = 1$; $d_{C,z_2^*} = \sqrt{(1-6'5)^2+(2-0'5)^2} = 5'7$

$d_{E,z_1^*} = \sqrt{(2-1)^2+(4-3)^2} = 0'707$; $d_{E,z_2^*} = \sqrt{(2-6'5)^2+(4-0'5)^2} = 5'7$

$d_{B,z_1^*} = \sqrt{(6-1)^2+(1-3)^2} = 5'38$; $d_{B,z_2^*} = \sqrt{(6-6'5)^2+(1-0'5)^2} = 0'707$

$d_{D,z_1^*} = \sqrt{(7-1)^2+(0-3)^2} = 6'7$; $d_{D,z_2^*} = \sqrt{(7-6'5)^2+(0-0'5)^2} = 0'707$

Lösung: POSITION: $z_1^*[1,3]$ $z_2^*[6'5, 0'5]$
GRUPPEN: [A, C, E] [B, D]

Beispiel: Kosten Umsatz Qualität #Rekla DLZ ....

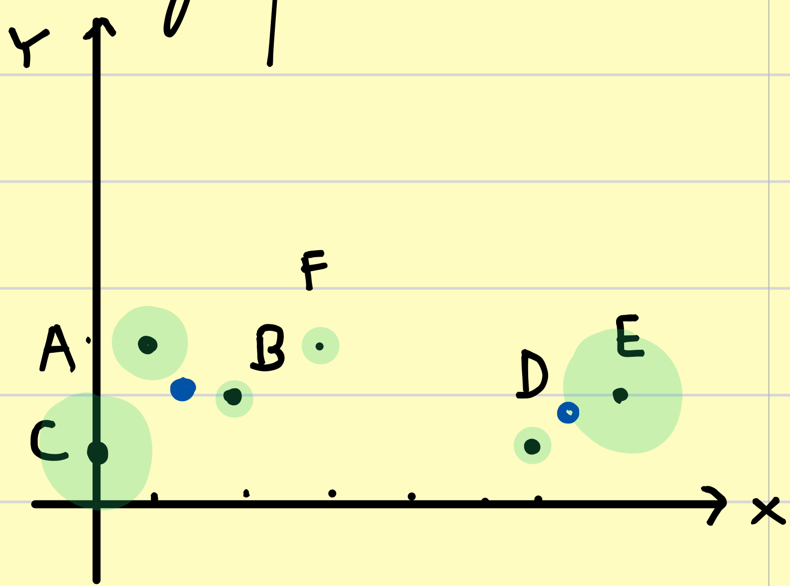
1
⋮
30000 Kunden 30.000

10 Gruppen ... python ... 1 Linie ...

Beispiel Übung: Die Positionen von 6 Werken mit unterschiedlichen Bedarfen an Rohware sind durch Ihre Koordinaten auf der Karte bestimmt.
Jedes Werk wird von einem der 2 geplanten Lägern

beliefert. Um die Fahrtkosten zu minimieren sollten die Läger so positioniert werden, dass sowohl die Werke möglichst nah sind, als auch die Bedarfe berücksichtigt werden. Bitte sprechen Sie der GF eine Empfehlung aus für die Lagerpositionen.

$$\begin{array}{ll} X & [1, 2, 0, 6, 7, 3] \\ Y & [3, 2, 1, 1, 2, 3] \\ \text{Bedarfe} & [2, 1, 3, 1, 3, 1] \\ & \;\; A \;\; B \;\; C \;\; D \;\; E \;\; F \end{array}$$



Anfangsgruppen: 1 [A, C, D] 2 [B, E, F]

HINWEIS: Gewichteter Mittelwert.

$$z_1 = \left[\frac{1 \cdot 2 + 0 \cdot 3 + 6 \cdot 1}{2+3+1}, \frac{3 \cdot 2 + 1 \cdot 3 + 1 \cdot 1}{2+3+1}\right] = [\ldots, \ldots]$$