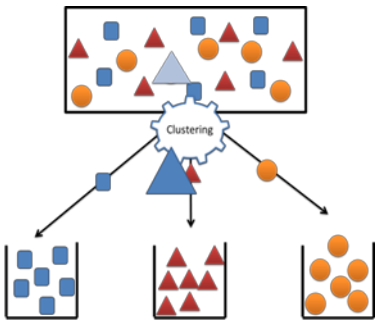
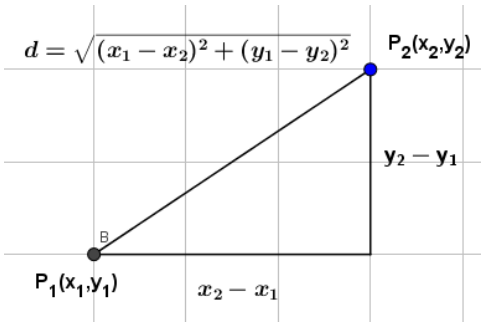
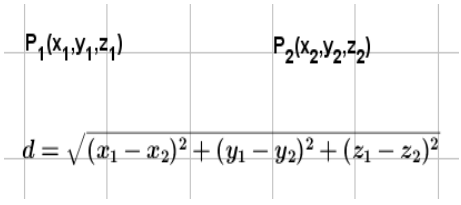



*Figuur 3.4 Het classificeren van objecten*




*Figuur 3.5 Stelling van Pythagoras*




*Figuur 3.6 Euclidische afstand bij drie dimensies*

# 3.3 Clusteranalyse

## Inleiding

Clusteranalyse is het classificeren of het groeperen in 'clusters' of 'klassen' van objecten op grond van hun kenmerken. Het doel van clusteranalyse is het vormen van deelverzamelingen die elk hun eigen gedeelde kenmerken bevatten.
Het doel van clusteranalyse is het verdelen van een dataset in groepen. Deze groepen en het aantal groepen zijn vooraf niet bekend. Het streven is zoveel mogelijk gelijkenis binnen een groep en zoveel mogelijk verschil tussen de groepen te krijgen. In tegenstelling tot classificatie: daar weten we de indeling in groepen al, en willen we een nieuw object in de juiste groep krijgen. Clusteranalyse is vooral bekend uit marktonderzoek en wordt veel ingezet om het koopgedrag van klanten te onderzoeken. Het doel van de analyse is niet het voorspellen van dit koopgedrag, maar het zoeken naar een beperkt aantal groepen klanten met hetzelfde koopgedrag. Online kan deze techniek heel goed worden ingezet om websitebezoek te onderzoeken. Hiermee kunnen verschillende doelgroepen op een effectieve wijze benaderd worden. Een bedrijf krijgt zo zicht op welk product of dienst een groep klanten het beste aansluit, en welke product eventueel niet haalbaar of minder bereikbaar zijn voor de klantengroep.
In de biologie zijn er meerdere gebieden waar clusteranalyse wordt toegepast. Denk bijvoorbeeld aan de classificatie van verschillende organismen. Elk organisme hoort bij een soort. Soorten kunnen op hun beurt weer worden onderverdeeld in lagere taxa, zoals ondersoort en variëteit. Soorten zelf worden samengevoegd in geslachten en deze weer in families en in taxa van nog hogere rang. Een ander voorbeeld van het gebruiken clustertechnieken in de biologie is het maken van groepen met genen die zie een bepaalde erfelijke ziekte kunnen bevatten. Door het gebruik van clustermethodes kunnen de groepen met genen gevonden worden. Als deze groepen bekend zijn, wordt het gemakkelijk een medicijn te ontwikkelen dat de erfelijke ziekte kan voorkomen of genezen.

## Afstandsmaten

**Afstand tussen twee punten**
In de wiskunde kan de afstand worden berekend als de wortel uit de som van de kwadraten van de verschillen tussen de coördinaten, volgens de stelling van Pythagoras, zie Figuur 3.5. In drie dimensies geldt analoog hiervoor de Euclidische afstand, zie Figuur 3.6

**Afstand tussen datapunten (objecten):**
Dit is een maat die aangeeft hoe groot de ‘overeenkomst’ of het ‘verschil’ is tussen twee kenmerk datapunten.
Er zijn twee bekende methoden om de afstand tussen waarnemingen:

De definitie van de **Euclidische afstand** is:

$$d(A,B) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2}$$

$$= \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}$$

Waarbij:

| $\boldsymbol{A}$ | $\boldsymbol{B}$ | $\boldsymbol{A-B}$ | $\boldsymbol{(A-B)^2}$ |
|---|---|---|---|
| $\boldsymbol{a_1}$ | $b_1$ | $a_1 - b_1$ | $(a_1 - b_1)^2$ |
| $\boldsymbol{a_2}$ | $b_2$ | $a_2 - b_2$ | $(a_2 - b_2)^2$ |
| | | | |
| $\boldsymbol{a_i}$ | $b_i$ | $a_i - b_i$ | $(a_i - b_i)^2$ |
| | | | |
| $\boldsymbol{a_n}$ | $b_n$ | $a_n - b_n$ | $(a_n - b_n)^2$ |

De definitie van de **Manhattan of City-block afstand** is

$$d(A,B) = \sum_{i=1}^{n}|a_i - b_i|$$

$$= |a_1 - b_1| + |a_2 - b_2| + \cdots + |a_n - b_n|$$

Waarbij:

| $\boldsymbol{A}$ | $\boldsymbol{B}$ | $\boldsymbol{A-B}$ | $\boldsymbol{\lvert A-B\rvert}$ |
|---|---|---|---|
| $\boldsymbol{a_1}$ | $b_1$ | $a_1 - b_1$ | $\lvert a_1 - b_1\rvert$ |
| $\boldsymbol{a_2}$ | $b_2$ | $a_2 - b_2$ | $\lvert a_2 - b_2\rvert$ |
| | | | |
| $\boldsymbol{a_i}$ | $b_i$ | $a_i - b_i$ | $\lvert a_2 - b_2\rvert$ |
| | | | |
| $\boldsymbol{a_n}$ | $b_n$ | $a_n - b_n$ | $\lvert a_n - b_n\rvert$ |

**Voorbeeld A**

Geef de afstand tussen de volgende objecten:

| | | | | | |
|---|---|---|---|---|---|
| Object $A$ | 10 | 12 | 15 | 13 | 9 |
| Object $B$ | 18 | 23 | 13 | 15 | 17 |

De Manhattan of City-block afstand is:

$d(A,B) =$

$= |10 - 8| + |12 - 23| + |15 - 13| + |13 - 15| + |9 - 17| = 25$

De Euclidische afstand is:

$d(A,B) =$

$$= \sqrt{\begin{matrix}(10 - 18)^2 + (12 - 23)^2 + (15 - 13)^2 + \\ (13 - 15)^2 + (9 - 17)^2\end{matrix}}$$

$= 16{,}16$

**Afstandsmatrix**

Een matrix van afstanden is een matrix waarvan de elementen de afstanden tussen de punten aangeven.

Twee dimensionale Euclidische afstand:

$d(a,b) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}$

**Voorbeeld B**

Gegeven een dataset met vier objecten. Geef de afstandsmatrix.

| | $x$ | $y$ |
|---|---|---|
| $A$ | 2 | 1 |
| $B$ | 4 | 2 |
| $C$ | 6 | 1 |
| $D$ | 7 | 2 |

De afstandsmatrix ziet eruit als volgt:

| Afstand | $A$ | $B$ | $C$ | $D$ |
|---|---|---|---|---|
| $A$ | 0 | 2.24 | 4 | 5.1 |
| $B$ | 2.24 | 0 | 2.24 | 3 |
| $C$ | 4 | 2,24 | 0 | 1.41 |
| $D$ | 5.1 | 5 | 1.41 | 0 |

De afstandsmaat die hier wordt gebruikt is de gewone Euclidische afstand.

## Afstand tussen clusters

Als je de afstand tussen elk paar van objecten weet, wat is dan de afstand tussen twee clusters $C_1$ en $C_2$?

**Single Linkage**

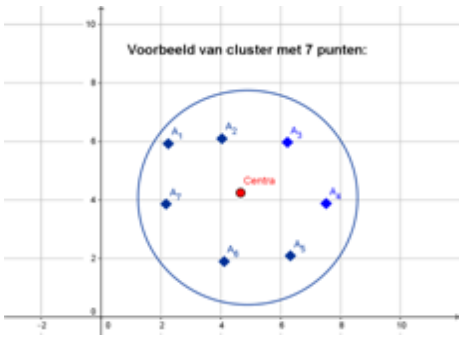
De afstand tussen twee clusters $C_1$ en $C_2$ is:

$d(C_1, C_2) = \min\{d(A,B) | A \epsilon C_1 \; en \;\; B \epsilon C_2\}$

**Euclidische afstand**

$$d(A,B) = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2}$$

$$= \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}$$

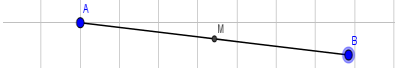



*Figuur 3.7 Het centrum van een cluster met zeven objecten*

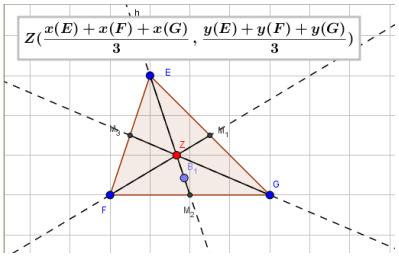
**Het centrum van een cluster**

In de wiskunde is het centrum van een cluster met twee punten $A$ en $B$ het **middelpunt** $M$ van het lijnstuk $[A, B]$.

$$M(\frac{x(A)+x(B))}{2}, \frac{y(A)+y(B)}{2})$$




Heb je te maken met drie punten, dan gebruik je het zwaartepunt. Hieronder zie je de drie zwaartelijnen in een driehoek die door één punt gaan. Dat punt heet het zwaartepunt $Z$ van de driehoek.




Het centrum van een cluster met drie punten $E$, $F$ en $G$ moet gezien worden als **het zwaartepunt** van een driehoek $\Delta EFG$

**Voorbeeld C**

Gegeven een dataset met zes punten.

| | $x$ | $y$ |
|---|---|---|
| $A_1$ | 1 | 3 |
| $A_2$ | 1 | 4 |
| $A_3$ | 2 | 2 |
| $B_1$ | 5 | 1 |
| $B_2$ | 5 | 2 |
| $B_3$ | 7 | 2 |

De clusters zijn $C_1 = \{A_1, A_2, A_3\}$ en $C_2 = \{B_1, B_2, B_3\}$.
De afstandsmatrix is

| | $B_1$ | $B_2$ | $B_3$ |
|---|---|---|---|
| $A_1$ | 4,47 | 4,12 | 6,03 |
| $A_2$ | 5 | 4,47 | 6,32 |
| $A_3$ | 3,16 | 3 | 5 |

Geef de minimale afstand tussen $C_1$ en $C_2$.

$\mathrm{d}(C_1, C_2) = \ d(C_1, C_2) = \min\{d(A, B) | A \in \mathrm{C}_1, \mathrm{B} \in \mathrm{C}_2\}$
$\mathrm{d}(A_3, B_2) = \sqrt{(5-2)^2 + (2-2)^2} = 3$
Met de groene kleur wordt de minimale afstand aangegeven.

Het centrum $C(x, y)$ van een cluster bepaal je door gemiddelde van alle punten $A_i(x_i, y_i \ldots\ldots z_i)$ in die cluster te

Om het **centrum** van een cluster te berekenen gebruik je de volgende formule

$$M(x, y, \ldots, z) = (\frac{\sum_1^n x_i}{n}, \frac{\sum_1^n y_i}{n}, \ldots, \frac{\sum_1^n z_i}{n})$$

Ofwel

$$M(x, y, \ldots, z) = (\frac{x_1+x_2+\cdots\ldots+x_n}{n}, \frac{y_1+y_2+\cdots\ldots+y_n}{n}, \ldots, \frac{z_1+z_2+\cdots\ldots+z_n}{n})$$

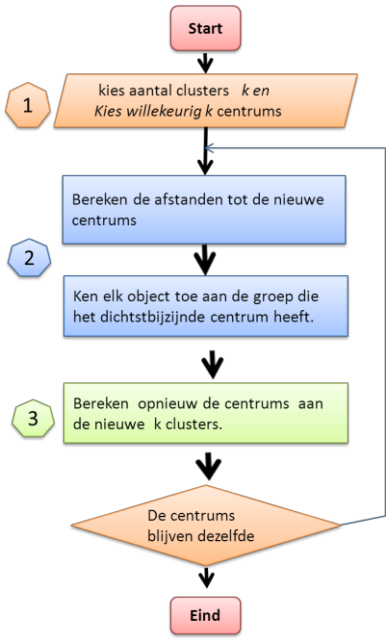
**Voorbeeld D**

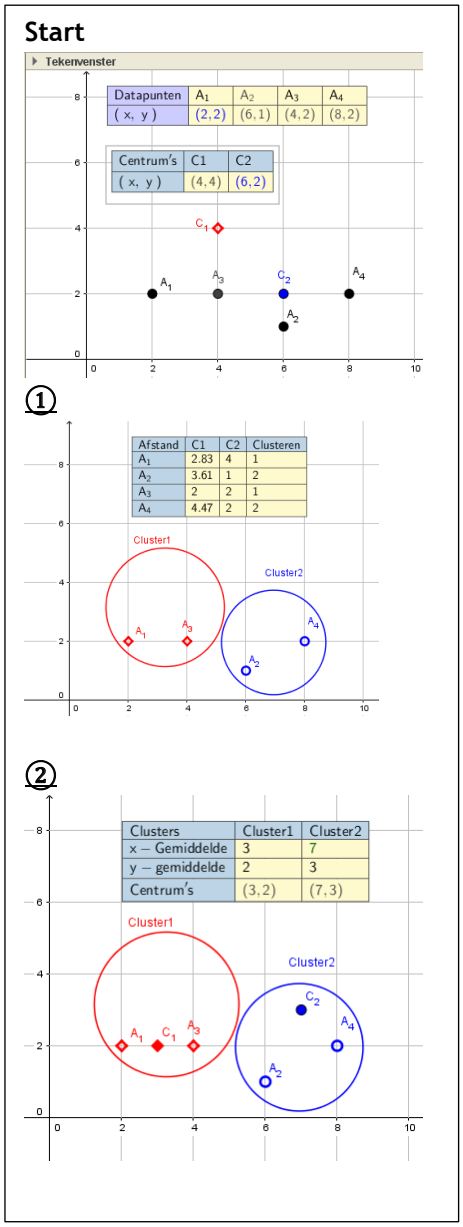
Gegeven een cluster $C$ die de volgende objecten bevat:

| Datapunten | $x$ | $y$ |
|---|---|---|
| $A_1$ | 0 | 0 |
| $A_2$ | 7 | 8 |
| $A_3$ | 4 | 8 |
| $A_4$ | 3 | 0 |

Bereken het centrum van de cluster $C$.

$C = \{A_1, A_2, A_3, A_4\}$
$M\ ((0+7+4+3)/4, (0+8+8+0)/4)$
Het centrum van de cluster is $M(3.5, 4)$.




# Clustermethoden

Er zijn twee methodes om tot clusters te komen: hiërarchische of partitiemethode. In deze paragraaf gaan we ons beperken tot niet-hiërarchisch clustering.
Niet-hiërarchisch of partitioneren betekent het stap voor stap verbeteren van een bestaande clustering.

Het basisalgoritme is hier:

a. Begin met een willekeurige clustering in k clusters. (nadeel: $k$ ligt vooraf vast)
b. Herhaal (totdat je meent klaar te zijn): stop een object in een andere cluster, zodanig dat de kwaliteit van de clusters verbetert

**K-means-clustering**

Het oudste en meest bekende clusteralgoritme is de K-means methode. Later zijn er veel varianten ontwikkeld die gebaseerd zijn op deze methode. $K$-means is een eenvoudige, iteratieve manier van clusteren. Vooraf wordt bepaald hoeveel clusters je wilt hebben. Om de optimale verbetering van een bestaande clustering ga je als volgt te werk:

**Start**: Kies de centra van de clusters eerste keer gewoon willekeurig.
① Daarna bepaal je van elk datapunt de afstand tot ieder centrum, en wijs het datapunt toe aan de cluster waarvan het centrum het dichtstbij is.
② Nadat alle datapunten aan een cluster zijn toegevoegd, bereken je de centrums van de clusters opnieuw. Hiervoor gebruik je de volgende formule:

$$C(x,y,,z) = \left(\frac{\sum_1^n x_i}{n}, \frac{\sum_1^n y_i}{n}, \ldots, \frac{\sum_1^n z_i}{n}\right)$$

③ Als de centrums niet veranderen, herhaal dan de stappen ① en ②. Dit gaat door de centrums niet meer veranderen, of totdat er een vooraf gekozen aantal stappen is geweest.

**Voorbeeld E**

Gegeven de volgende dataset:

| Datapunten | $x$ | $y$ |
|---|---|---|
| $A_1$ | 2 | 2 |
| $A_2$ | 4 | 2 |
| $A_3$ | 6 | 1 |
| $A_4$ | 8 | 2 |

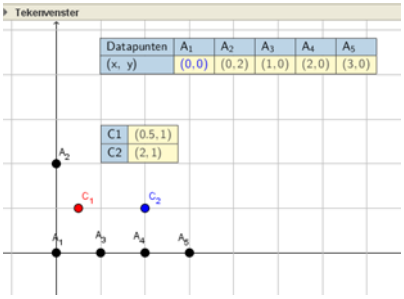
We bepalen vooraf dat we twee clusters willen, dus $k = 2$. Geef de uitwerking van de eerste iteratie.

**Start**

Kies twee willekeurig centrums, bijvoorbeeld:

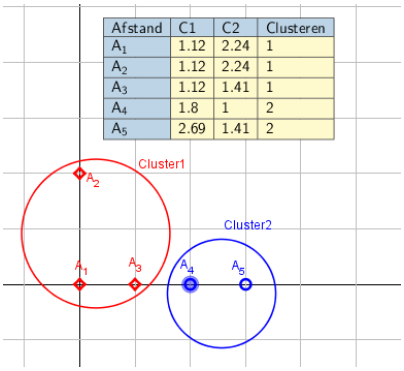
| $M_1 = (0.5,1)$ |
|---|

**Start**

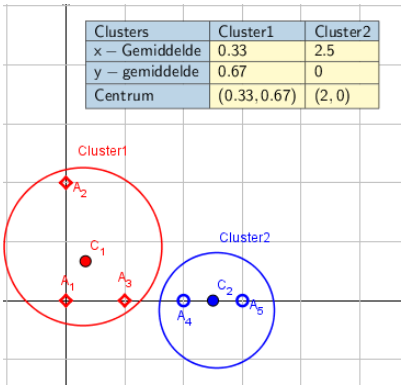



**Iteratie1**

①

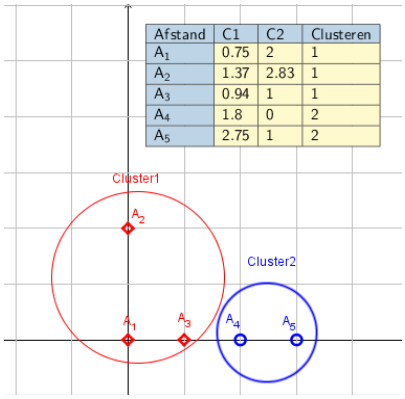



②




**Iteratie 2**

①




| |
|---|
| $M_2 = (2,1)$ |

**Iteratie 1**

①

Als afstandsmaat gebruik je de normale (Euclidische) afstand.

| Afstand | $C_1$ | $C_2$ | Toewijzing |
|---|---|---|---|
| $A_1$ | 2,83 | 4 | 1 |
| $A_2$ | 3,61 | 1 | 2 |
| $A_3$ | 2 | 2 | 1 |
| $A_4$ | 4,47 | 2 | 2 |

De nieuwe clusters zijn:

| | |
|---|---|
| $C_1$ | $\{A_1, A_3\}$ |
| $C_2$ | $\{A_2, A_4\}$ |

②

De nieuwe centrums zijn:

| |
|---|
| $M_1 = (3,2)$ |
| $M_2 = (7,3)$ |

**Voorbeeld F**

Gegeven de volgende dataset:

| Datapunten | $x$ | $y$ |
|---|---|---|
| $A_1$ | 0 | 0 |
| $A_2$ | 2 | 0 |
| $A_3$ | 0 | 1 |
| $A_4$ | 0 | 2 |
| $A_5$ | 3 | 0 |

**Start**

We zijn op zoek naar twee clusters, dus $k = 2$. Als afstandsmaat kun je de normale (Euclidische) afstand gebruiken. Geef de uitwerking van de eerste twee iteraties.

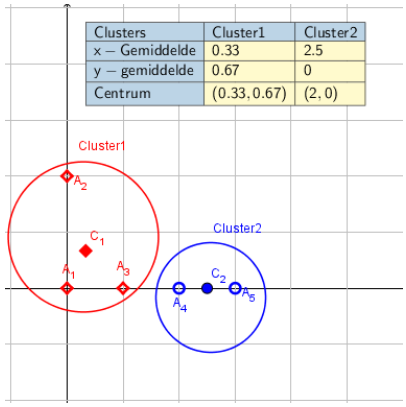
| |
|---|
| $M_1 = (0.5,1)$ |
| $M_2 = (2,1)$ |

**Iteratie 1**

①

| Afstand | $C_1$ | $C_2$ | Clusteren |
|---|---|---|---|
| $A_1$ | 1,12 | 2,24 | 1 |
| $A_2$ | 1,12 | 2,24 | 1 |
| $A_3$ | 1,12 | 1,41 | 1 |
| $A_4$ | 1,8 | 1 | 2 |
| $A_5$ | 2,69 | 1,41 | 2 |

②




Nieuwe clusters:

| | |
|---|---|
| $C_1$ | $\{A_1, A_2, A_3\}$ |
| $C_2$ | $\{A_4, A_5\}$ |

②

De nieuwe centrums zijn

| |
|---|
| $M_1 = \left(\frac{1}{3}, \frac{2}{3}\right)$ |
| $M_2 = (2,0)$ |

③

Herhaal nu de stappen ① en ② ⟶ **Iteratie 2**

**Iteratie 2**

①

| Afstand | $C_1$ | $C_2$ | Toekenning |
|---|---|---|---|
| $A_1$ | 0,75 | 2 | 1 |
| $A_2$ | 1,37 | 2,83 | 1 |
| $A_3$ | 0,94 | 1 | 1 |
| $A_4$ | 1,8 | 0 | 2 |
| $A_5$ | 2,75 | 1 | 2 |

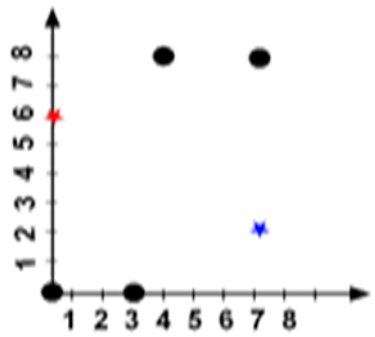
②

Er niets veranderd. De centrums blijven dezelfde.

$M_1 = \left(\frac{1}{3}, \frac{2}{3}\right)$

$M_2 = (2,0)$

In dit geval kunt je zeggen dat $k$-means algoritme convergeert.




*Figuur 3.10 De objecten van opdracht 4*

# Vragen en opdrachten

1.

Gegeven is een dataset met 6 datapunten.

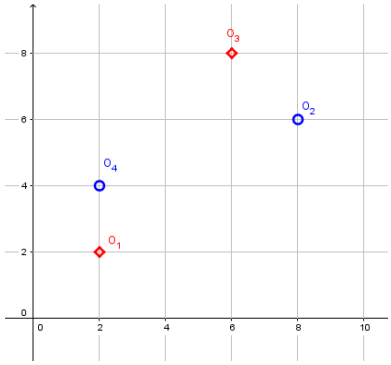
| Data | $X_1$ | $X_2$ | $X_3$ | $X_4$ |
|---|---|---|---|---|
| $A_1$ | 6 | 3 | 4 | 5 |
| $A_2$ | 2 | 3 | 5 | 4 |
| $A_3$ | 5 | 4 | 6 | 3 |
| $A_4$ | 9 | 1 | 1 | 8 |
| $A_5$ | 8 | 2 | 0 | 9 |
| $A_6$ | 8 | 0 | 1 | 8 |

Als afstandsmaat gebruiken we de normale (Euclidische) afstand
Bereken de Euclidische afstand van $A_1$ en $A_2$

Vul de volgende tabel in:

| Afstand | $A_4$ | $A_5$ | $A_6$ |
|---|---|---|---|
| $A_1$ | 5,57 | 6,08 | 5,57 |
| $A_2$ | 9,22 | 9,33 | 8,77 |
| $A_3$ | 8,66 | 9,22 | 8,66 |

Clusteren deze gegevens in twee clusters:
$C_1 = \{A_1, A_2, A_3\}$ en $C_2 = \{A_4, A_5, A_6\}$.
Geef de afstand tussen clusters $C_1$ en $C_2$.
Bereken de centrums van clusters $C_1$ en $C_2$.




*Figuur 3.9 De objecten in opdracht 2*

2.

Gegeven de volgende vier objecten:

| Object | $x_1$ | $x_2$ |
|---|---|---|
| $O_1$ | 2 | 2 |
| $O_2$ | 8 | 6 |
| $O_3$ | 6 | 8 |
| $O_4$ | 2 | 4 |

We willen deze gegevens clusteren in twee clusters ($k = 2$), met behulp van het k-means algoritme.
Initialiseren het algoritme: objecten 1 en 3 in één cluster $C_1$ en objecten 2 en 4 in de andere cluster $C_2$.

Noteer $C_1 = \{O_1, O_3\}$ en $C_2 = \{O_2, O_4\}$
Als afstandsmaat gebruik je de normale (Euclidische) afstand.
Bereken het centrum $M_1$ van $C_1$ en het centrum $M_2$ van $C_2$
Vul de onderstaande tabel in en bereken daarmee de afstand tussen de cluster $C_1$ en $C_2$.

| Afstand | $O_2$ | $O_4$ |
|---|---|---|
| $O_1$ | | |
| $O_3$ | | |

Bereken de afstanden tot de centrums $M_1$ en $M_2$ en vul de volgende tabel in:

| Afstand | $M_1$ | $M_2$ |
|---|---|---|
| $O_1$ | | |
| $O_2$ | | |
| $O_3$ | | |
| $O_4$ | | |

3.
Gegeven is een dataset met $6$ datapunten (zie tabel hieronder):

| Data | $x_1$ | $x_2$ |
|---|---|---|
| $D_1$ | 6 | 3 |
| $D_2$ | 2 | 3 |
| $D_3$ | 5 | 4 |
| $D_4$ | 9 | 1 |
| $D_5$ | 8 | 2 |
| $D_6$ | 8 | 0 |

We willen deze gegevens clusteren in twee clusters ($k = 3$) met behulp van het $k$-means algoritme.
We willen objecten $D_1$ en $D_2$ in één cluster $C_1$, de objecten $D_3$ en $D_4$ in de cluster $C_2$ en $D_5$ en $D_6$ in de cluster $C_2$.

Dus: $C_1 = \{D_1, D_2\}$, $C_2 = \{D_3, D_4\}$ en $C_3 = \{D_5, D_6\}$

Als afstandsmaat gebruik je de normale (euclidische) afstand
Bereken de centrums $M_1$ voor $C_1$, $M_2$ voor $C_2$ en $M_3$ voor $C_3$.
Bereken de afstanden tot de centrums $M_1$ en $M_2$ en vul de volgende tabel in:

| Afstand | $M_1$ | $M_2$ | $M_3$ |
|---|---|---|---|
| $O_1$ | | | |
| $O_2$ | | | |
| $O_3$ | | | |
| $O_4$ | | | |

4.
Gegeven is een dataset met vier objecten.

| Objecten | $x$ | $y$ |
|---|---|---|
| $A_1$ | 0 | 0 |
| $A_2$ | 7 | 8 |
| $A_3$ | 4 | 8 |
| $A_4$ | 3 | 0 |

Cluster de objecten door middel van het $k$-means algoritme. Geef de uitwerking van de tot twee iteraties.
Start het algoritme met de willekeurige centrums $(0,6)$ en $(7,2)$
Noteer voor elke iteratie welke cluster gevormd worden en wat de centrums van de cluster zijn en vul deze resultaten in de onderstaande tabel in:

| | Clusters | Centrums |
|---|---|---|
| Start | | $M_1 = (0,6)$<br>$M_2 = (7,2)$ |
| Iteratie 1 | $C_1 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$<br>$C_2 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$ | $M_1 =$<br>$M_2 =$ |
| Iteratie 2 | $C_1 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$<br>$C_2 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$ | $M_1 =$<br>$M_2 =$ |

5.

Een reisagentschap wil zijn klanten opdelen in twee clusters ($k = 2$), op basis van leeftijd en duur van geboekte vakanties. De informatie van het reisagentschap is gegeven in de onderstaande tabel. De leeftijd van klanten uitgedrukt in aantal jaren, de duur van de vakantie in het aantal dagen.

| ID | Leeftijd | Duur vakantie |
|---|---|---|
| $TO_1$ | 19 | 3 |
| $TO_2$ | 25 | 8 |
| $TO_3$ | 43 | 14 |
| $TO_4$ | 61 | 14 |
| $TO_5$ | 30 | 7 |
| $TO_6$ | 22 | 10 |

Cluster de objecten door middel van het $k$-means algoritme

Kies $TO_1$ als centrum voor cluster $C_1$ en Kies $TO_5$ als centrum voor cluster $C_2$.

Noteer voor elke iteratie welk cluster gevormd wordt en wat de centrums zijn. Vul deze resultaten in de onderstaande tabel in:

| Iteratie | Clusters | Centrums |
|---|---|---|
| Start | | $M_1 = (19,3)$<br>$M_2 = (30,7)$ |
| Iteratie 1 | $C_1 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$<br>$C_2 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$ | $M_1 =$<br>$M_2 =$ |
| Iteratie 2 | $C_1 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$<br>$C_2 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$ | $M_1 =$<br>$M_2 =$ |

6.

Hieronder staan gegevens van $4$ (fictieve) studenten van een data mining cursus. We noteren respectievelijk het aantal bijgewoonde lessen, het aantal dagen examenvoorbereiding, en of ze al dan niet voor het examen kwamen opdagen:

$D = \{S_1(0.0, 0, 1), S_2(7.8\ ,2, 1), S_3(4.8, 2, 1), S_4(3.0, 0, 1)\}$

We willen deze objecten clusteren, wat we kunnen doen door middel van de $k$-means methode.

Kies $k = 2$, en als initiële willekeurige cluster centrums: $M_1 = (0.6,\ 1,\ 0)$ en $M_2 = (7.2, 1, 0)$.

Pas de $k$-means methode toe op $D$ tot een maximum van $3$ stappen. Noteer voor elke iteratie welke clusters gevormd worden en wat de centrums zijn.

Convergeert de methode? Zo ja, leg uit.
Vul de resultaten in de onderstaande tabel in:

| | Clusters | Centrums |
|---|---|---|
| Start | | $M_1 = (0{,}6{,}1{,}0)$<br>$M_2 = (7, 2, 1, 0)$ |
| Iteratie 1 | $C_1 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$<br>$C_2 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$ | $M_1 =$<br>$M_2 =$ |
| Iteratie 2 | $C_1 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$<br>$C_2 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$ | $M_1 =$<br>$M_2 =$ |
| Iteratie 3 | $C_1 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$<br>$C_2 = \{\ldots\ldots\ldots\ldots\ldots\ldots\}$ | $M_1 =$<br>$M_2 =$ |
| …… | …… | …… |