



Universiteit Leiden

Leiden Institute of Advanced Computer Science  
Research & Education

- **Introductie**
- **Multisets**
- **Visualisatie**
- **Voorbeeld**
- **Methode 1**
- **Nieuwe meth.**
- **Alignment**
- **Vragen?**

# Criminele carrieres vergelijken

Jeroen F. J. Laros en Walter A. Kusters  
LIACS  
Universiteit Leiden



- Introductie
- Multisets
- Visualisatie
- Voorbeeld
- Methode 1
- Nieuwe meth.
- Alignment
- Vragen?

## Meerdere misdaden in verschillende categorieën.

Om te kunnen werken met meerdere misdaden per tijdseenheid, hebben we het begrip **Multiset** nodig.



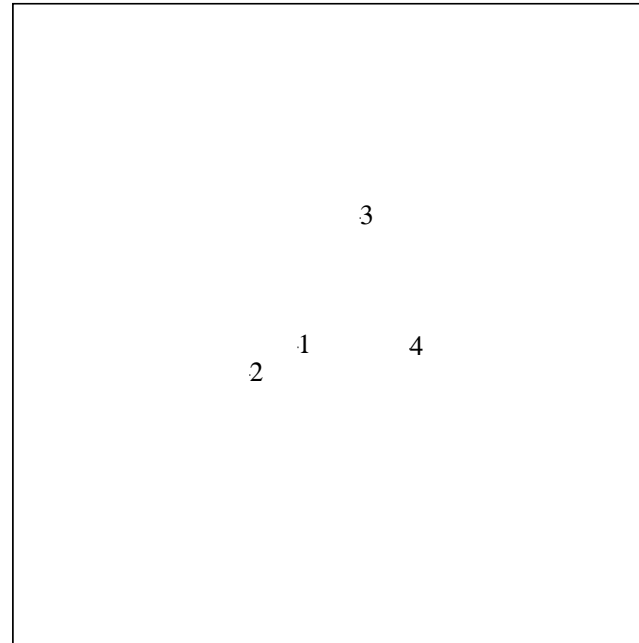
Multiset  $X$  met onderliggende set  $S(X)$ .

Voorheen (in de zgn. “market basket analysis”) werd er alleen met de onderliggende set gewerkt.



- Introductie
- Multisets
- Visualisatie
- Voorbeeld
- Methode 1
- Nieuwe meth.
- Alignment
- Vragen?

## Visualisatie

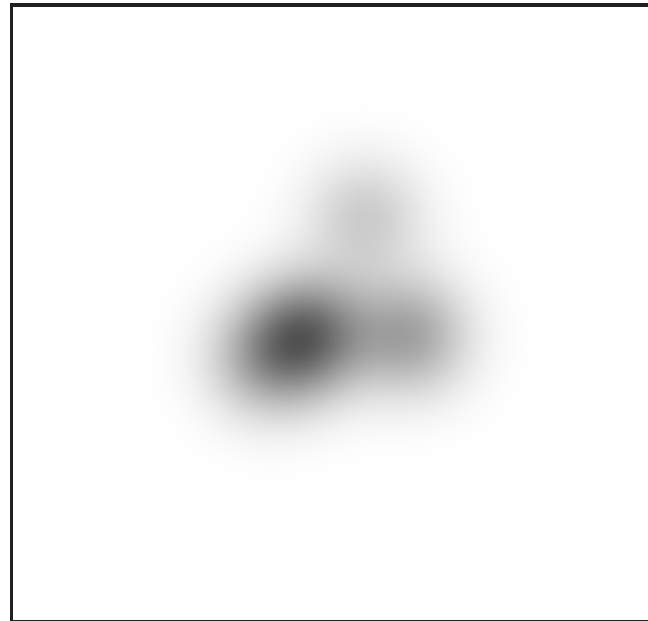


Hier zien we een clustering van vier punten. We gebruiken een **push and pull** algoritme om de punten zo goed mogelijk te plaatsen. In dit plaatje is er expres een fout gemaakt om de volgende slide interessant te maken.



- **Introductie**
- **Multisets**
- **Visualisatie**
- **Voorbeeld**
- **Methode 1**
- **Nieuwe meth.**
- **Alignment**
- **Vragen?**

## Fouten weergeven



Hier zien we een dieptekaart van de fout in het vorige plaatje. Dit kan een onderzoeker een goed idee geven van de betrouwbaarheid van de visualisatie en eventueel slechts delen van het originele plaatje gebruiken.



- Introductie
- Multisets
- Visualisatie
- Voorbeeld
- Methode 1
- Nieuwe meth.
- Alignment
- Vragen?

## Tien criminelen, vier categorieën

	1	2	3	4	5	6	7	8	9	10
$A$	0	2	10	0	0	0	0	2	0	2
$B$	0	0	0	2	0	0	2	4	0	2
$C$	0	0	0	0	1	0	2	0	3	2
$D$	0	0	0	0	0	1	1	0	5	2

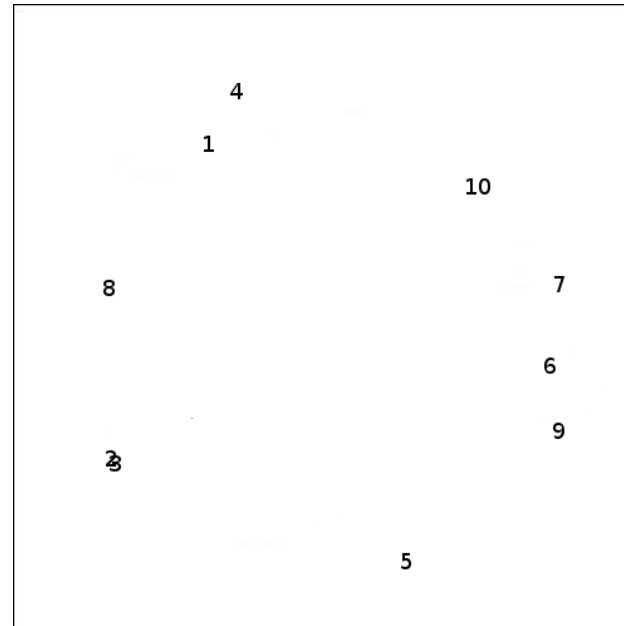
We hebben een dataset met vier categorieën,  $A$ ,  $B$ ,  $C$  en  $D$  van oplopende zwaarte. We hebben 10 criminelen waarvoor in elke categorie het aantal misdaden is genoteerd.

- 1: Onschuldig.
- 2: Incidentele lichte crimineel.
- 6: Eenmalige dader van een zware misdaad.
- 9: Zware crimineel.
- 10: All-rounder.



- Introductie
- Multisets
- Visualisatie
- Voorbeeld
- Methode 1
- Nieuwe meth.
- Alignment
- Vragen?

## De oude methode

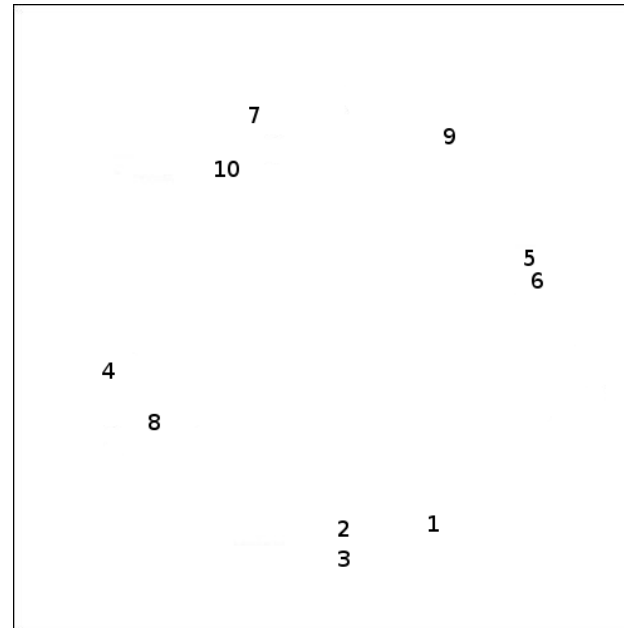


Merk op dat crimineel 2 en 3 als gelijk worden gezien. Ze zijn inderdaad veroordeeld voor dezelfde misdaden, maar het aantal misdaden verschilt behoorlijk.



- Introductie
- Multisets
- Visualisatie
- Voorbeeld
- Methode 1
- Nieuwe meth.
- Alignment
- Vragen?

## De nieuwe methode.

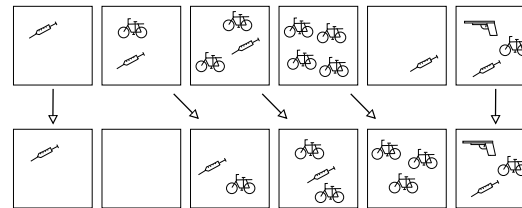


Met de nieuwe methode kunnen we ten eerste rekening houden met de hoeveelheid misdaden, maar ook met de zwaarte van een misdrijf. Zodoende kunnen we het plaatje hierboven construeren.



- Introductie
- Multisets
- Visualisatie
- Voorbeeld
- Methode 1
- Nieuwe meth.
- Alignment
- Vragen?

## Alignment



Met onze nieuwe afstandsmaat kunnen we niet alleen per jaar, maar ook de afstand voor hele carrieres berekenen met behulp van concepten uit de biologie.

Alignment zoekt naar een rangschikking van de elementen zodanig dat de volgorde behouden blijft en het aantal aanpassingen minimaal is.





- **Introductie**
- **Multisets**
- **Visualisatie**
- **Voorbeeld**
- **Methode 1**
- **Nieuwe meth.**
- **Alignment**
- **Vragen?**

## Vragen?

This research is part of the DALE (Data Assistance for Law Enforcement) project as financed in the ToKeN program from the Netherlands Organization for Scientific Research (NWO) under grant number 634.000.430.

