

# ODISSEI Community Conference

22 October 2019 | Galgenwaard, Utrecht



**ODISSEI**

Open Data Infrastructure for Social Science and Economic Innovations





# Measuring segregation using a Social Network of the Netherlands

Jan van der Laan <[dj.vanderlaan@cbs.nl](mailto:dj.vanderlaan@cbs.nl)>

ODISSEI Community Conference – 22 October 2019

# Measuring segregation

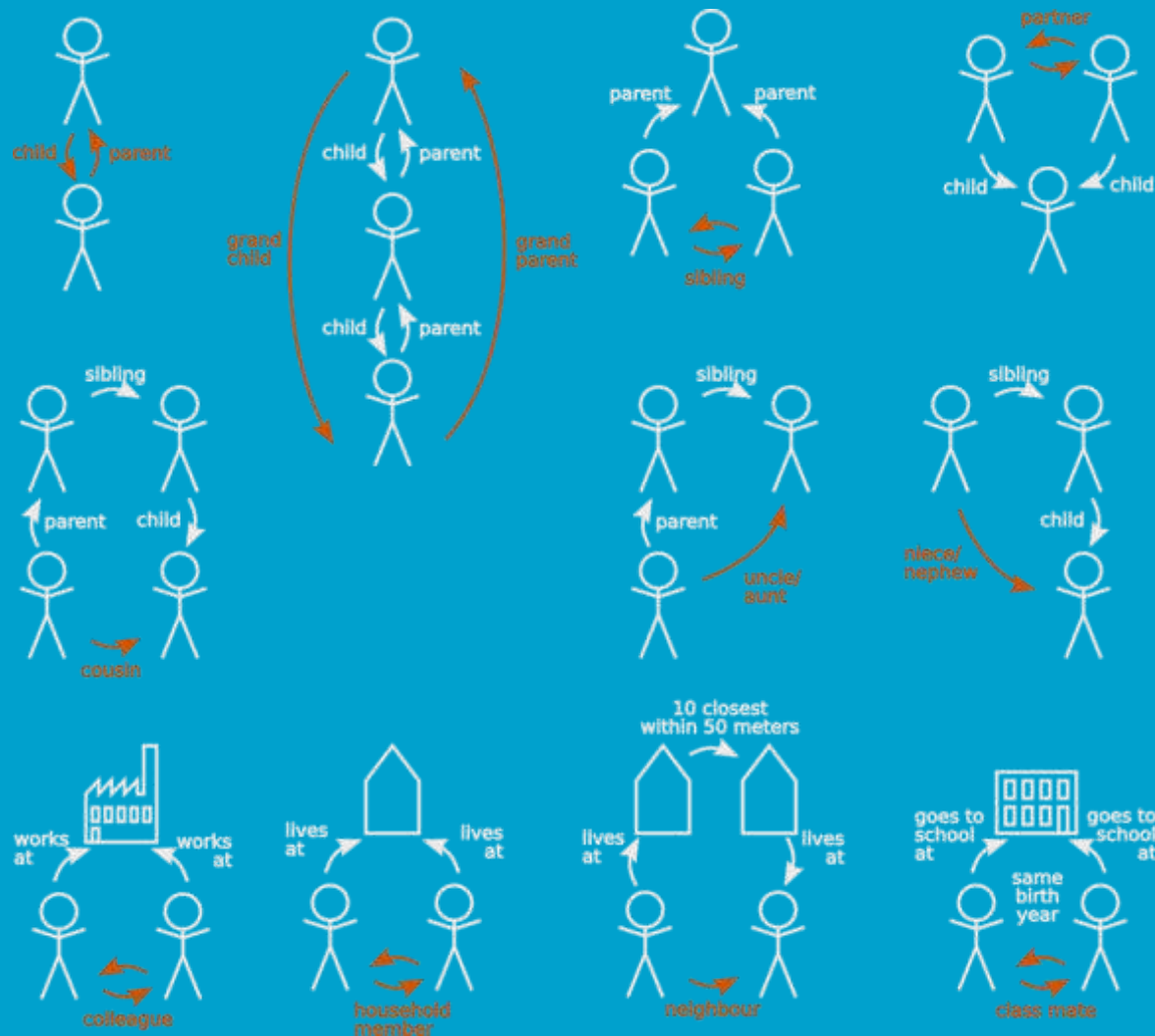
Segregation is an indicator for social inequality

Goal: measure segregation at the individual level

- Is determined by the behaviour of individuals
- Expect large differences between persons
- Easy to aggregate
- What determines that somebody is less/more segregated?
- What effect does segregation have on someone?

Measure using local **ego-networks** and not just spatial (as is traditional)





# Social Network of the Netherlands

Family  
Work  
Neighbours  
School

16.9M persons  
793M relations



# Exposure and isolation

**Exposure:** to what degree is a person exposed to persons from group  $h$ :

*$E_i(h)$  = fraction of persons from group  $h$  in ego-network of person  $i$*

**Isolation:** to what degree is person exposed to only members of their own group ( $g_i$ ):

$$I_i(h) = E_i(g_i)$$

Isolation is a measure of segregation



# Normalisation

When no segregation:

Expect that exposure is proportional to size of the group

$$\eta_i^2 = \frac{I_i - f_{g_i}}{1 - f_{g_i}}$$

$\eta_i^2 = 1$  Complete segregation

$\eta_i^2 > 0$  Segregation

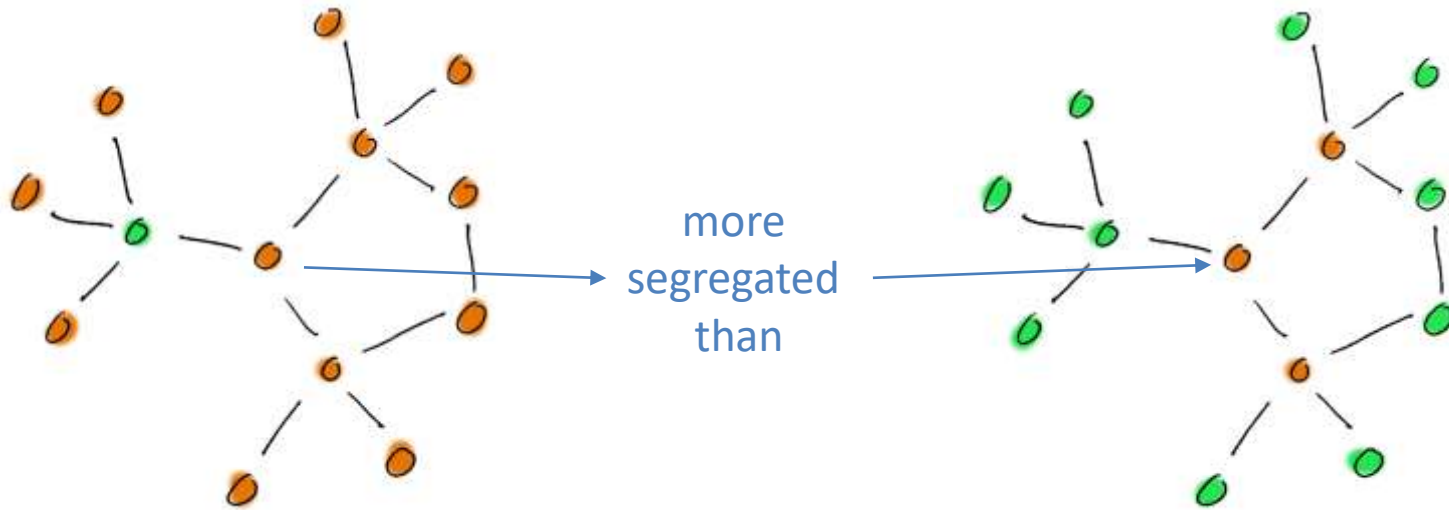
$\eta_i^2 = 0$  No segregation

$\eta_i^2 < 0$  *Heterophily*



# Ego-network of a person

Segregation not only depends on the direct contacts of a person, but also of the contacts of those contacts.

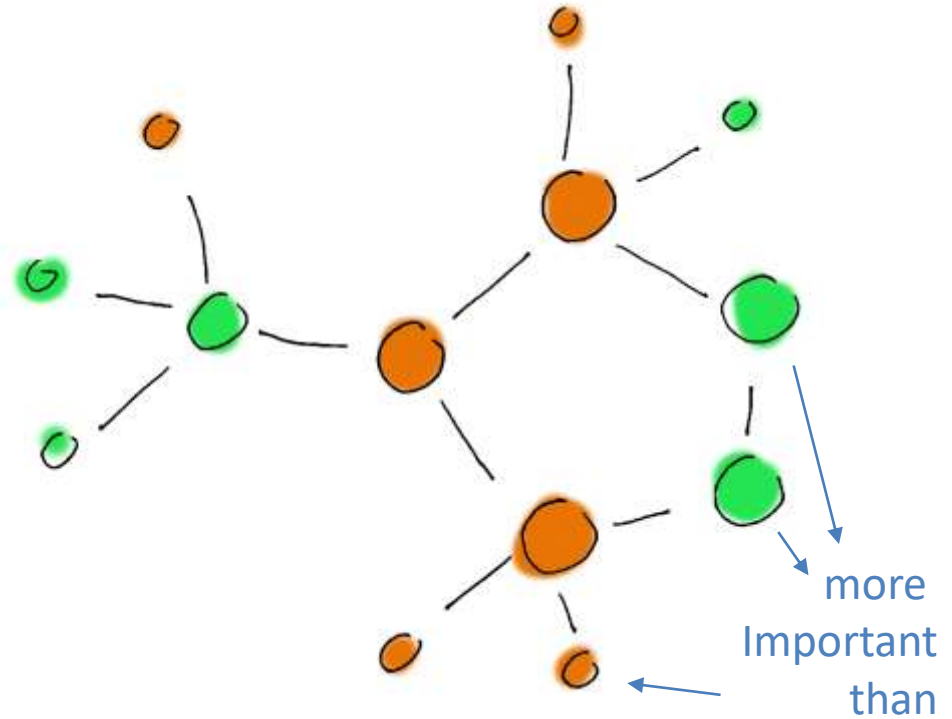




# Local random walk

Give persons further away (in the network) from a person a smaller weight.

Give important persons in the network a higher weight





# Calculations

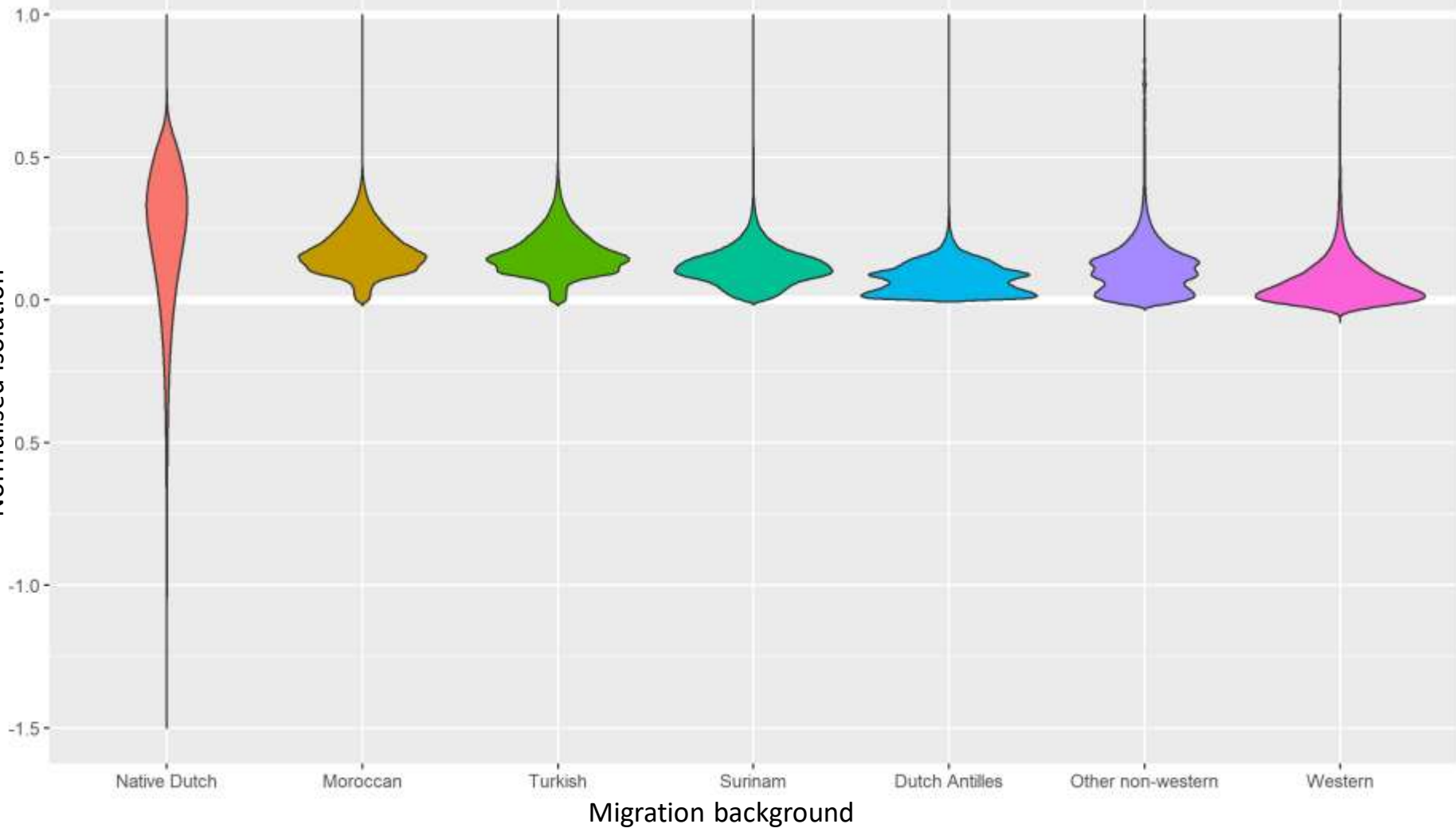
- Local random walk is computational intensive
- Network is large: 800 million relations
- Memory was main bottle neck: needed approx. 200 GB
- Sensitive data

## ODISSEI Secure Super Computer

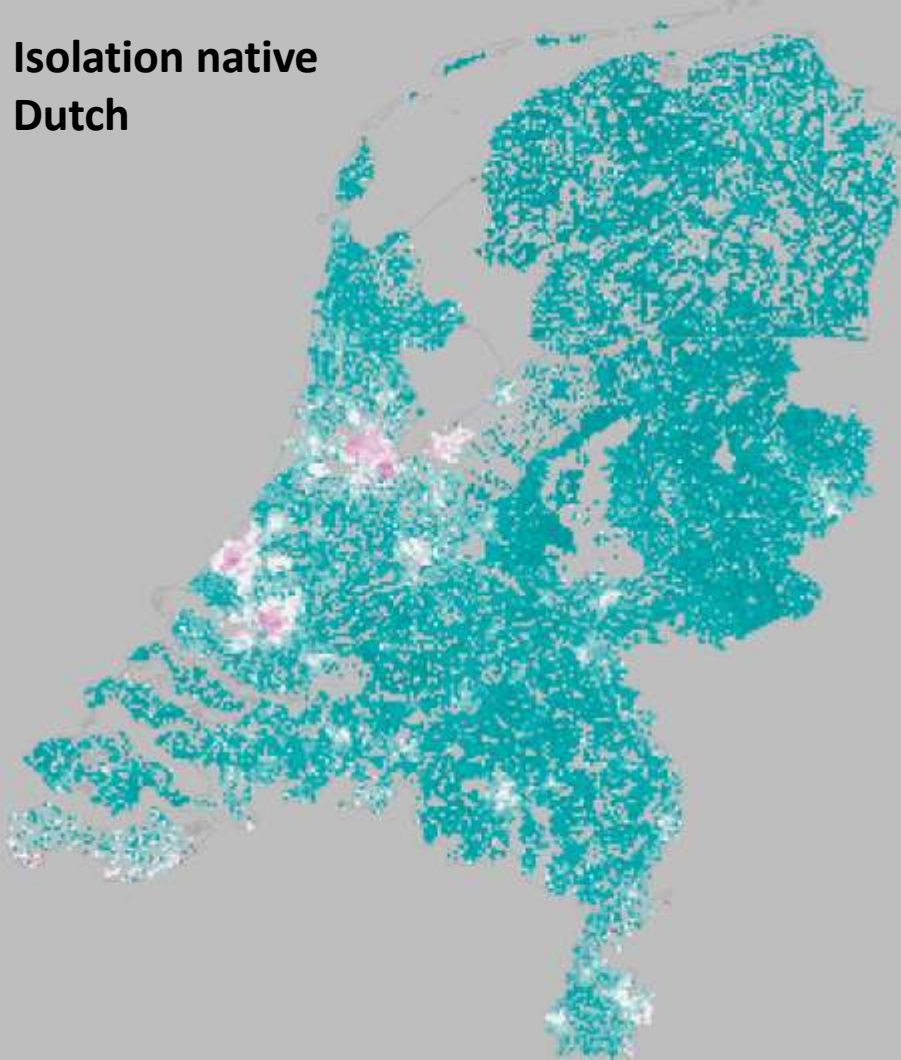
- Machine 250GB memory; multiple cores
- Was able to use existing (self-written) R programme



Normalised isolation



**Isolation native  
Dutch**



Complete segregation

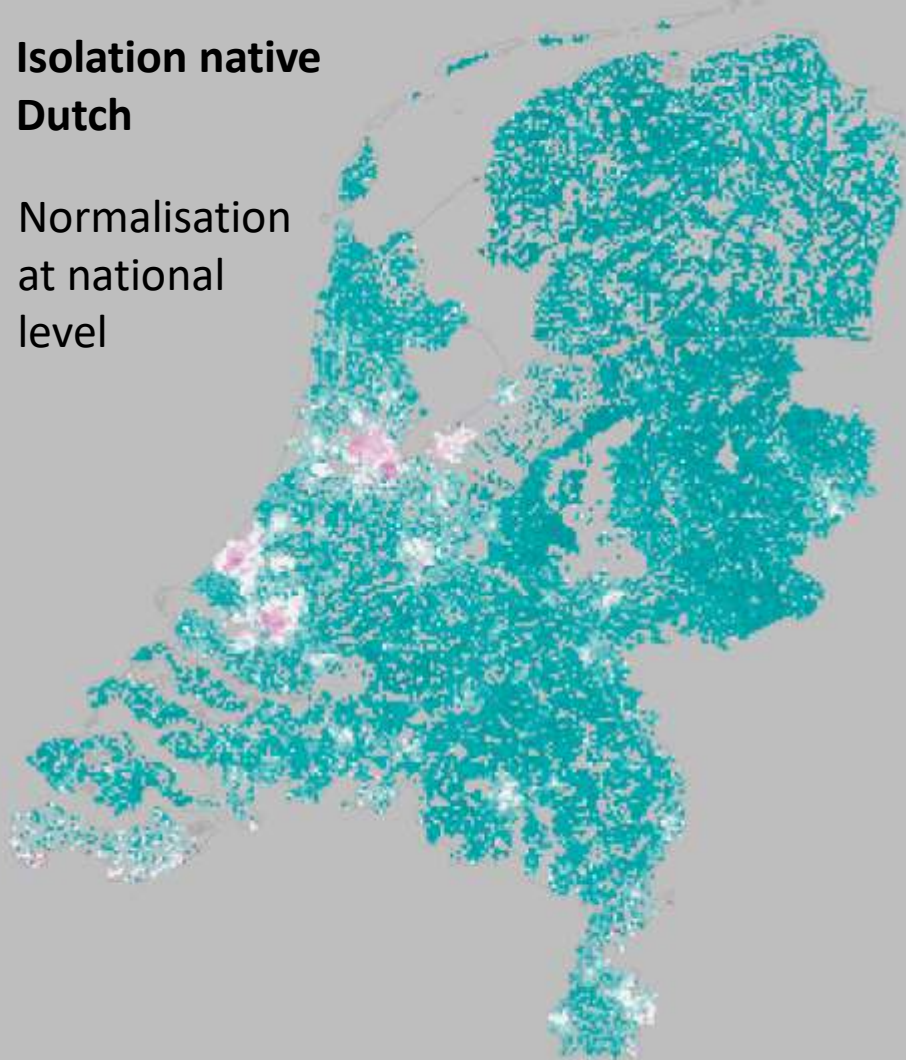
No segregation

Isolation is as expected

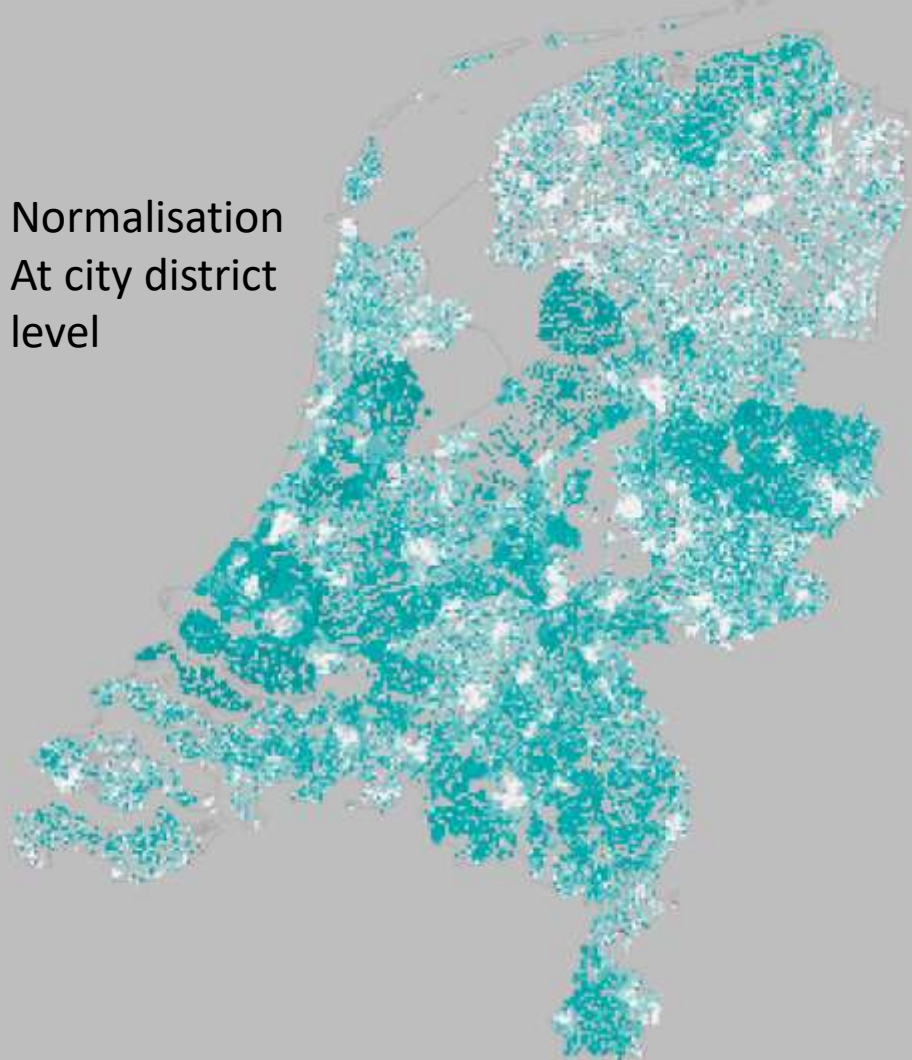
Isolation lower than expected

**Isolation native  
Dutch**

Normalisation  
at national  
level



Normalisation  
At city district  
level



# Exposure COROP Groot Amsterdam

Exposed to migration background



Migration background



# Conclusion

- It is now possible to calculate segregation at individual level
- Possible to compare persons to their ego-network

## Short term plans:

- Segregation on education level and income
- Effect of the different types of networks on segregation
- Derive temporal network



# ODISSEI Community Conference

22 October 2019 | Galgenwaard, Utrecht



**ODISSEI**

Open Data Infrastructure for Social Science and Economic Innovations