Done so far

- Read the paper: Social norms and obesity prevalence: From cohort to system dynamics models

- Data investigation

- Descriptive statistics

Interesting finding

- People of different education have on average the same amount of links
- Native people (autochtonen) have on average more links than other groups
- The connections of people with a higher education are higher than that of people with a lower education
- Age plays a role in the amount of connections a person has

Questions

- Do we need ID’s to make our network

- Do we need more data in order to make a network for Amsterdam

- Builds our research on that of : Social norms and obesity prevalence: From cohort to system dynamics models

Planning to do

- **Read literature on Stochastic Actor model** and Link imputation/Link prediction

- **Looking at the diversity of links (gender/etnc/education)**

- **Growth of connections over time (age)**

- **City of Amsterdam distributions and match with own distributions**
- Opinion dynamics
- Network literatur/multi layer network

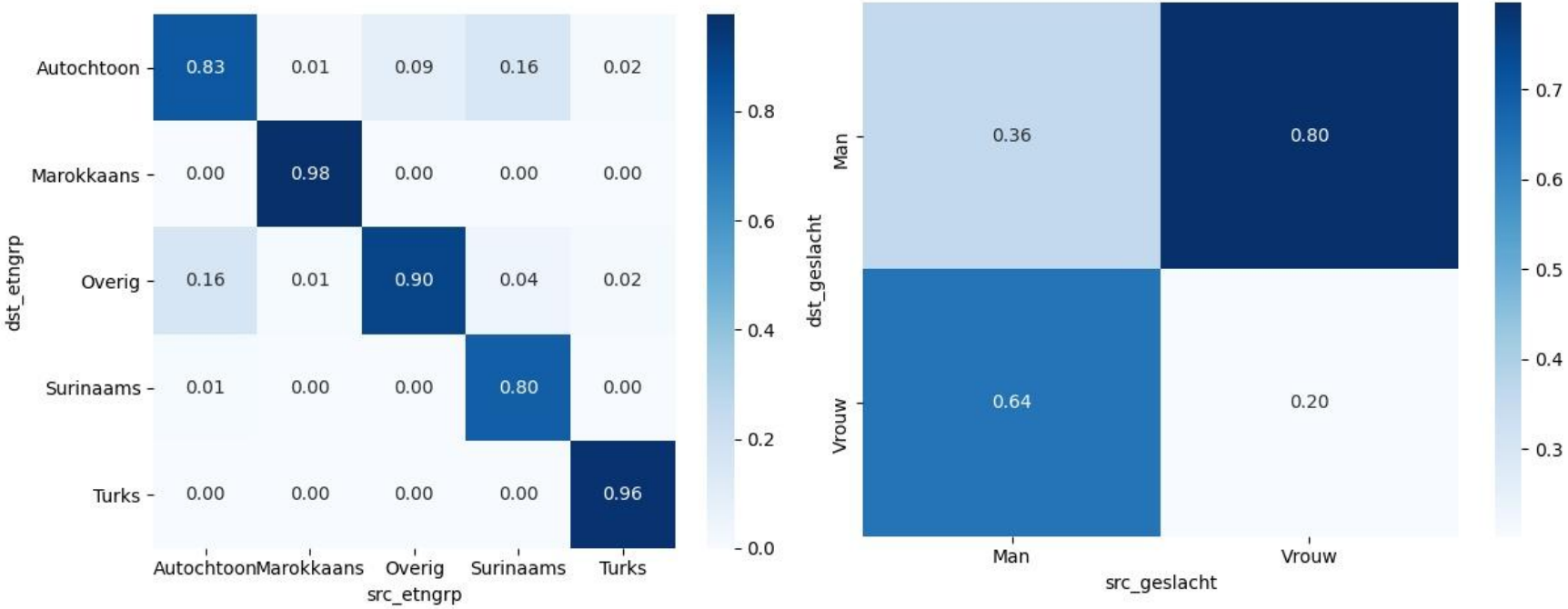
Done this week

- Made heatmaps for homophily + statistics (welch's t-test)

- SOAM
  - Measurements, implementation, possible problems

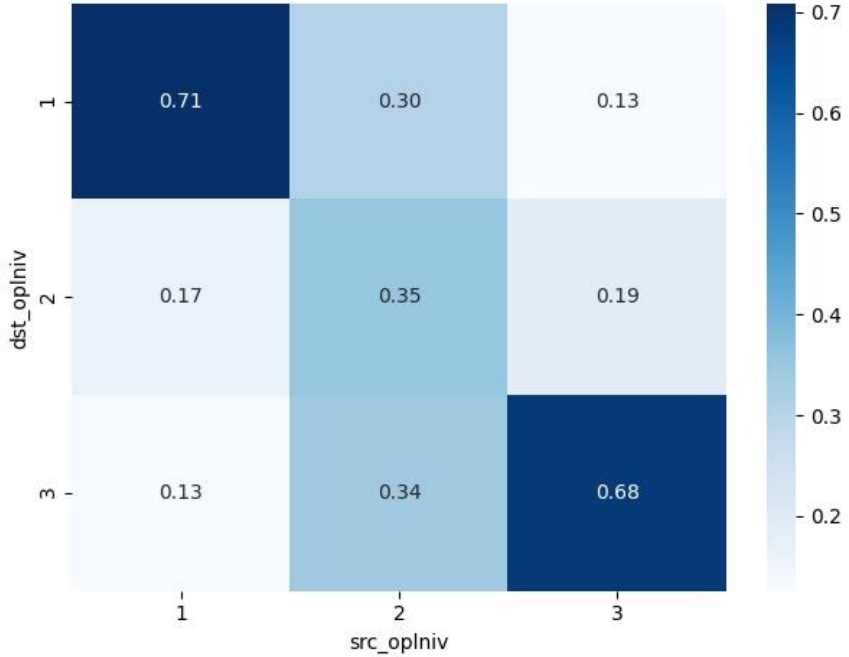
- Made a small first implementation of a network
  - To have a better understanding of how it can be implemented
  - To look if I need some extra data

Household

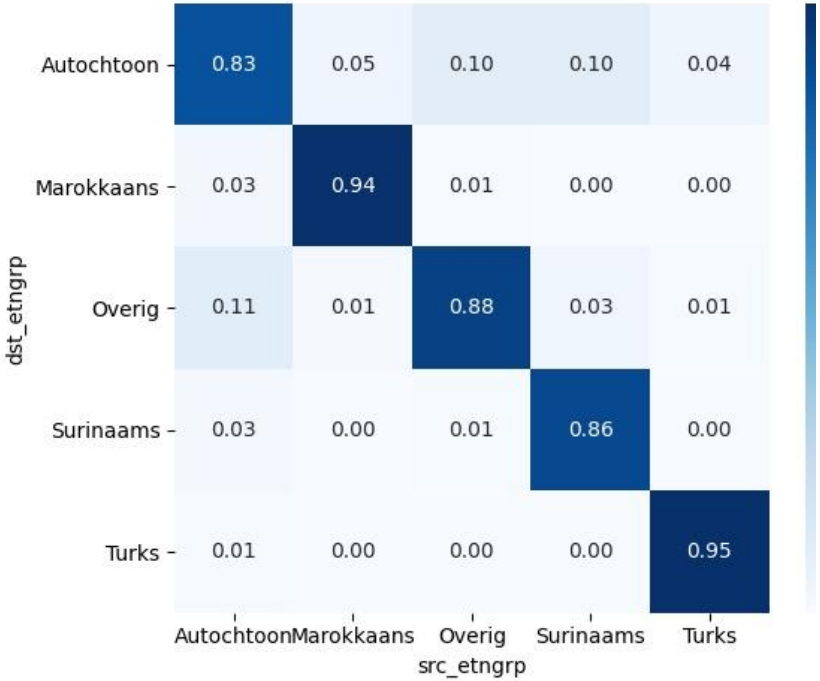

Autochtoon
Marokkaans
Overig
Surinaams
Turks
dst_etngrp
0.83 0.01 0.09 0.16 0.02
0.00 0.98 0.00 0.00 0.00
0.16 0.01 0.90 0.04 0.02
0.01 0.00 0.00 0.80 0.00
0.00 0.00 0.00 0.00 0.96
AutochtoonMarokkaans Overig Surinaams Turks
src_etngrp
0.8 0.6 0.4 0.2 0.0
Man
Vrouw
dst_geslacht
0.36 0.80
0.64 0.20
Man Vrouw
src_geslacht
0.7 0.6 0.5 0.4 0.3

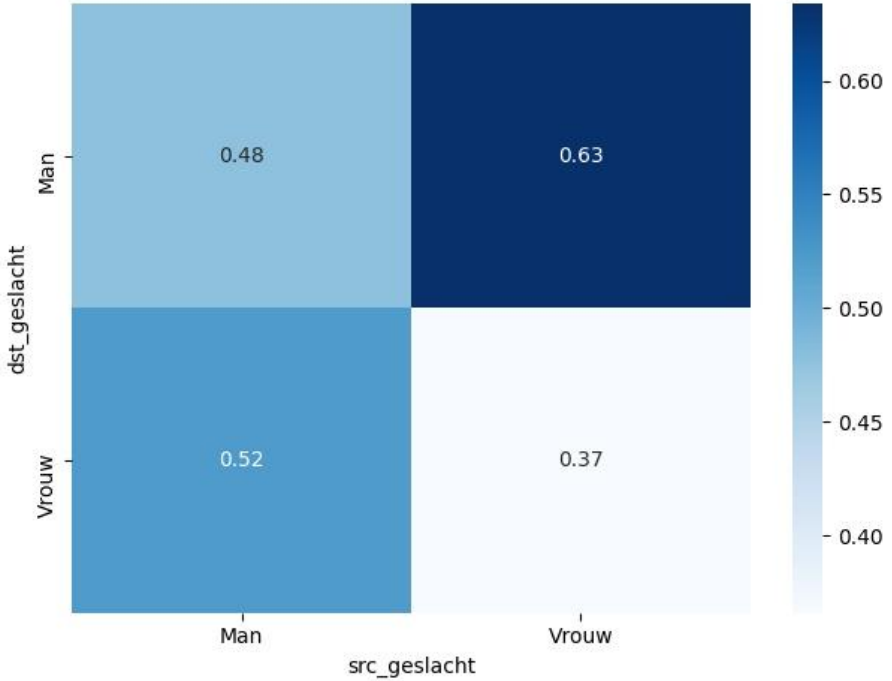
Household


dst_lft
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
[70,80)
[80,120]
src_lft
[0,20) [20,30) [30,40) [40,50) [50,60) [60,70) [70,80) [80,120]
0.65 0.07 0.06 0.25 0.22 0.01 0.00 0.00
0.03 0.29 0.13 0.07 0.05 0.03 0.00 0.00
0.08 0.47 0.74 0.20 0.02 0.01 0.00 0.00
0.11 0.07 0.06 0.31 0.14 0.01 0.00 0.00
0.12 0.06 0.01 0.17 0.45 0.14 0.01 0.00
0.01 0.03 0.00 0.01 0.11 0.55 0.14 0.00
0.00 0.00 0.00 0.00 0.01 0.24 0.73 0.28
0.00 0.00 0.00 0.00 0.00 0.00 0.12 0.72
0.7 0.6 0.5 0.4 0.3 0.2 0.1 0.0

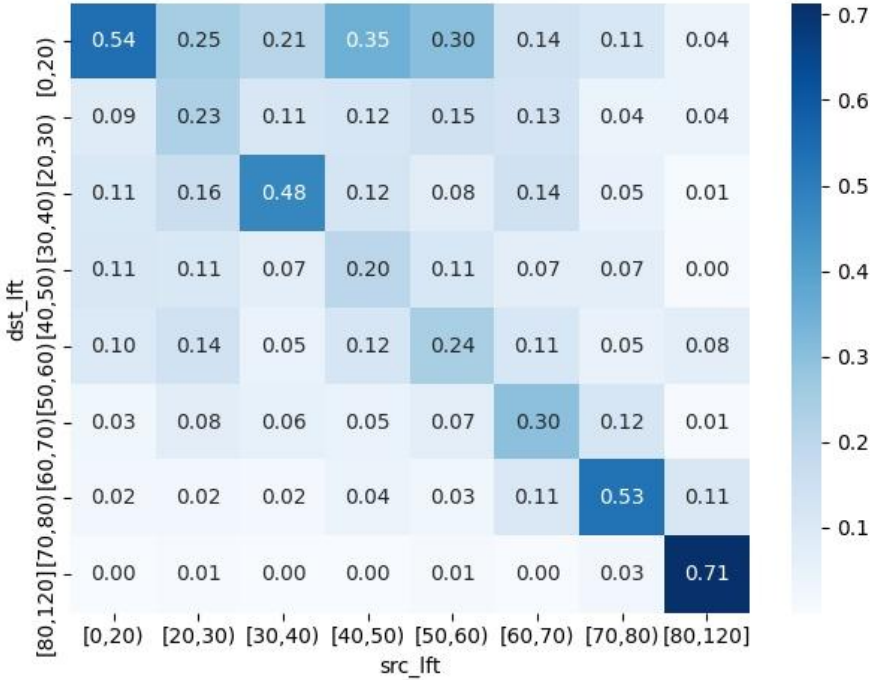

dst_oplniv
1
2
3
src_oplniv
1 2 3
0.71 0.30 0.13
0.17 0.35 0.19
0.13 0.34 0.68
0.7 0.6 0.5 0.4 0.3 0.2

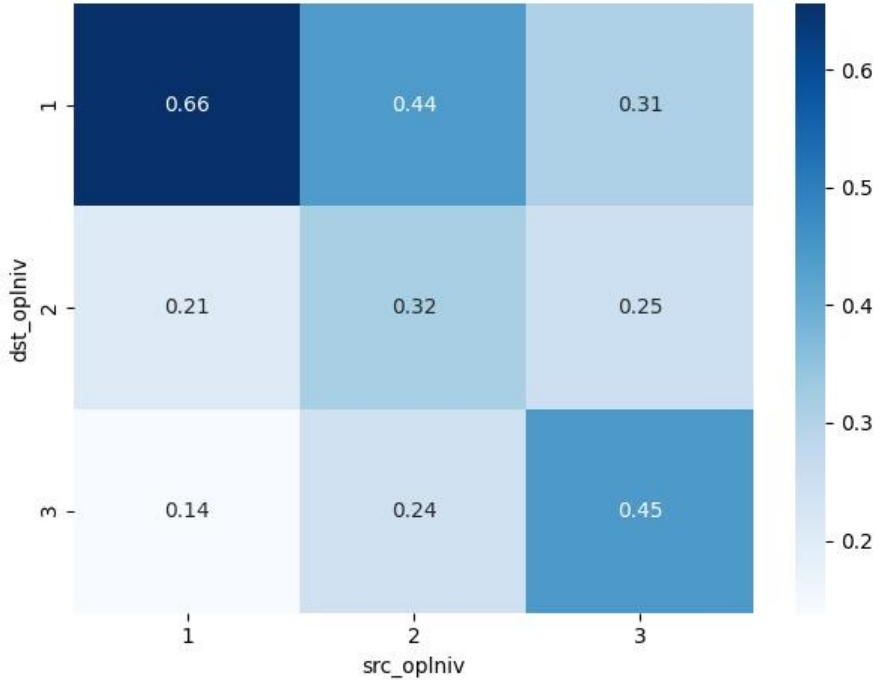
Family


Autochtoon
Marokkaans
Overig
Surinaams
Turks
dst_etngrp
0.83 0.05 0.10 0.10 0.04
0.03 0.94 0.01 0.00 0.00
0.11 0.01 0.88 0.03 0.01
0.03 0.00 0.01 0.86 0.00
0.01 0.00 0.00 0.00 0.95
Autochtoon
Marokkaans
Overig
Surinaams
Turks
src_etngrp

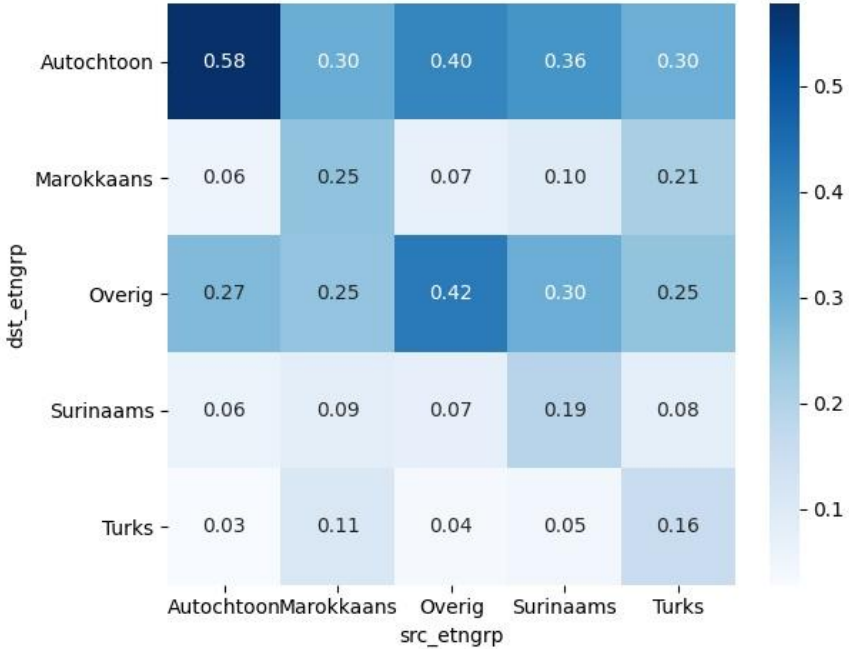

Man
Vrouw
dst_geslacht
0.48 0.63
0.52 0.37
Man
Vrouw
src_geslacht
0.60
0.55
0.50
0.45
0.40

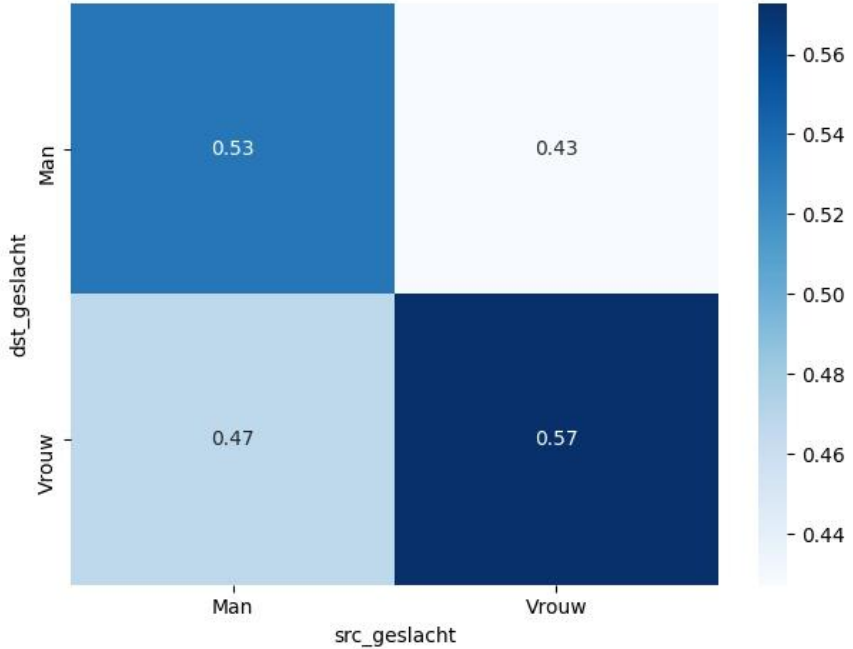
Family


dst_lft
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
[70,80)
[80,120]
0.54 0.25 0.21 0.35 0.30 0.14 0.11 0.04
0.09 0.23 0.11 0.12 0.15 0.13 0.04 0.04
0.11 0.16 0.48 0.12 0.08 0.14 0.05 0.01
0.11 0.11 0.07 0.20 0.11 0.07 0.07 0.00
0.10 0.14 0.05 0.12 0.24 0.11 0.05 0.08
0.03 0.08 0.06 0.05 0.07 0.30 0.12 0.01
0.02 0.02 0.02 0.04 0.03 0.11 0.53 0.11
0.00 0.01 0.00 0.00 0.01 0.00 0.03 0.71
[0,20) [20,30) [30,40) [40,50) [50,60) [60,70) [70,80) [80,120]
src_lft
0.7
0.6
0.5
0.4
0.3
0.2
0.1

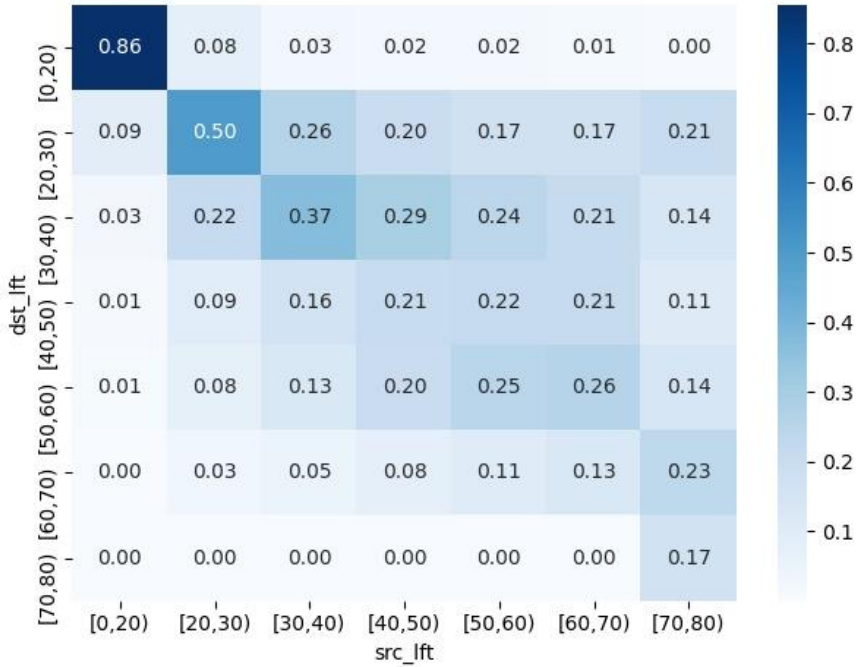

dst_oplniv
1
2
3
0.66 0.44 0.31
0.21 0.32 0.25
0.14 0.24 0.45
1 2 3
src_oplniv
0.6
0.5
0.4
0.3
0.2

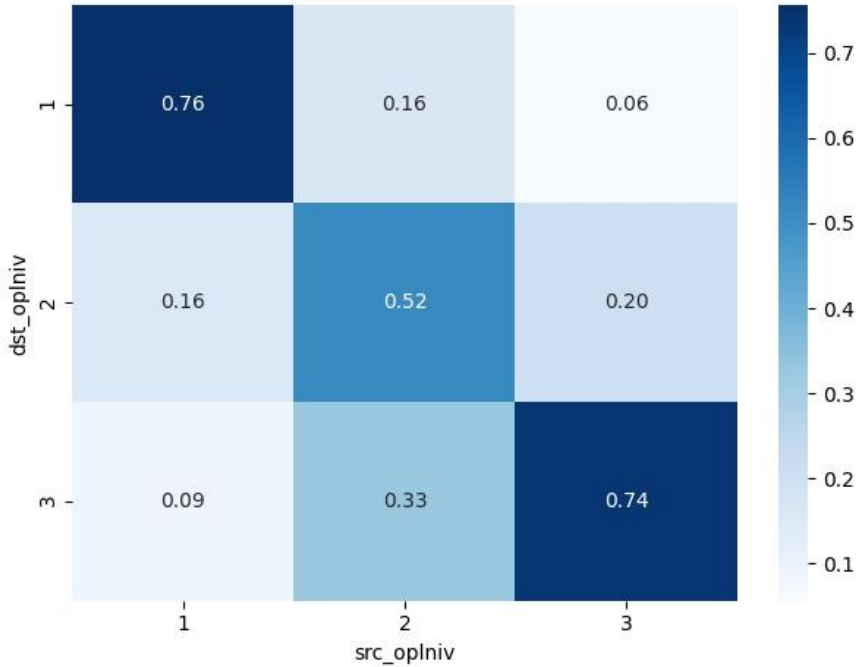
Work/School


Autochtoon
Marokkaans
Overig
Surinaams
Turks
dst_etngrp
0.58 0.30 0.40 0.36 0.30
0.06 0.25 0.07 0.10 0.21
0.27 0.25 0.42 0.30 0.25
0.06 0.09 0.07 0.19 0.08
0.03 0.11 0.04 0.05 0.16
AutochtoonMarokkaans Overig Surinaams Turks
src_etngrp
0.5
0.4
0.3
0.2
0.1

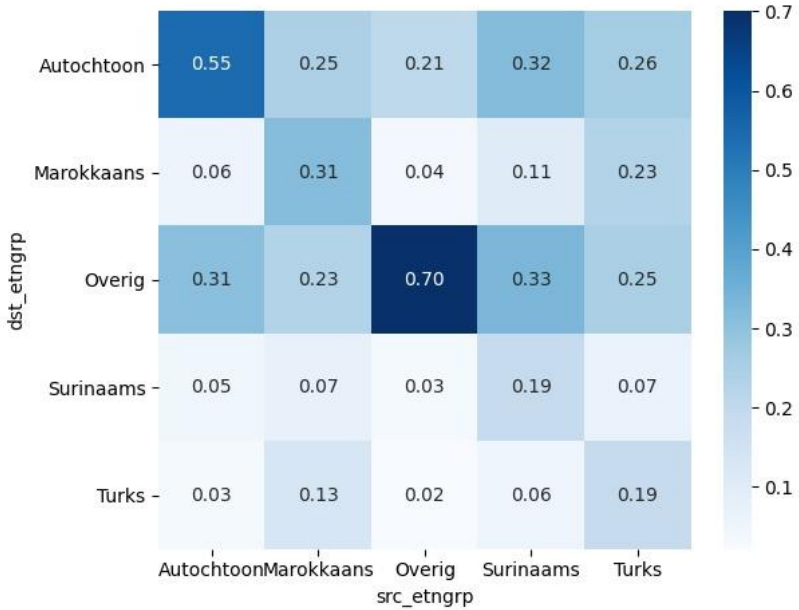

Man
Vrouw
dst_geslacht
0.53 0.43
0.47 0.57
Man Vrouw
src_geslacht
0.56
0.54
0.52
0.50
0.48
0.46
0.44

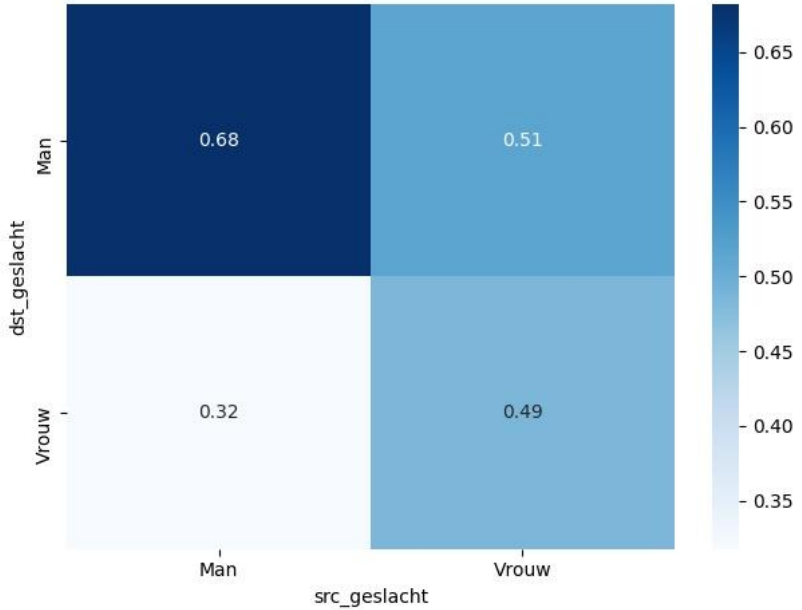
Work/School


dst_lft
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
[70,80)
0.86 0.08 0.03 0.02 0.02 0.01 0.00
0.09 0.50 0.26 0.20 0.17 0.17 0.21
0.03 0.22 0.37 0.29 0.24 0.21 0.14
0.01 0.09 0.16 0.21 0.22 0.21 0.11
0.01 0.08 0.13 0.20 0.25 0.26 0.14
0.00 0.03 0.05 0.08 0.11 0.13 0.23
0.00 0.00 0.00 0.00 0.00 0.00 0.17
[0,20) [20,30) [30,40) [40,50) [50,60) [60,70) [70,80)
src_lft
0.8 0.7 0.6 0.5 0.4 0.3 0.2 0.1

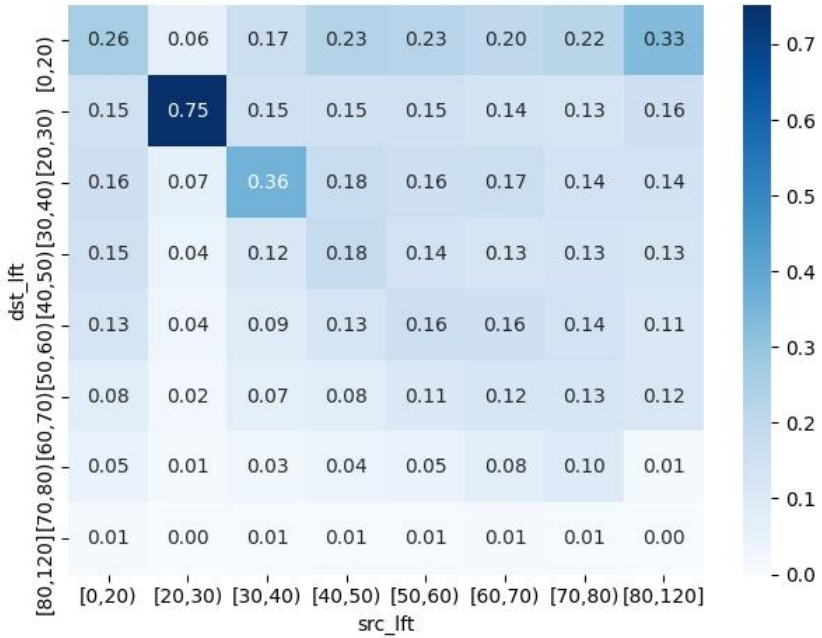

dst_oplniv
1
2
3
0.76 0.16 0.06
0.16 0.52 0.20
0.09 0.33 0.74
1 2 3
src_oplniv
0.7 0.6 0.5 0.4 0.3 0.2 0.1

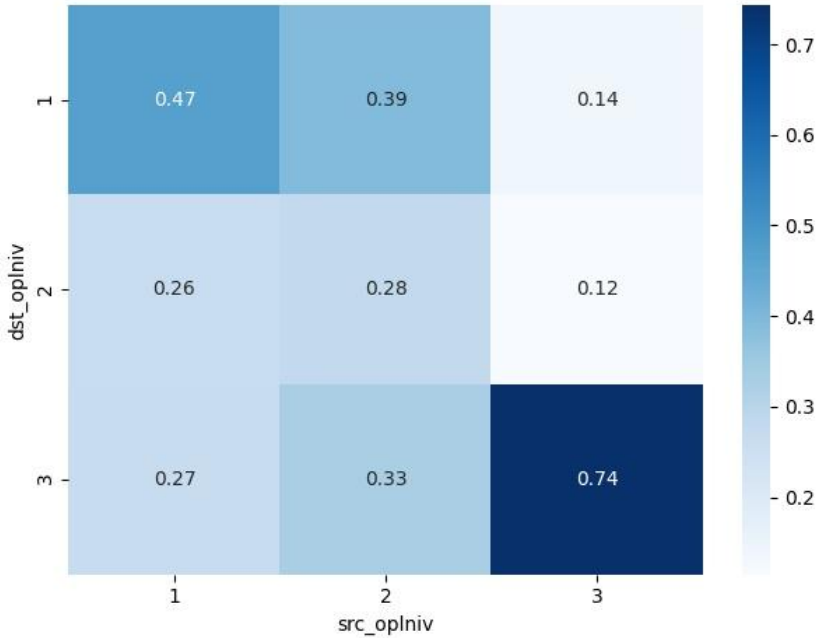
Neighborhood





Neighborhood





Interesting measurements

- Reciprocity
- Transitive triplets
- Out-degree
- In-degree

Problems with SOAM

- Formally the actors in the network are fully aware (. In practice, the actors only need more limited information)
- No more than a few hundred actors

Possible problems with the data

- Missing of the education level in the n_table
-

To do

- Connections work/school ⇒ based on?
- Theils
- Distribution of groups V
- Distribution of network V
- Use education from other tables to implement education in n_tab V
- Mutual Information
- https://regio-monitor.nl/ V

- Normalized index homophily ⇒ segregation V
- Multi layer literatur, link imputation
-

Done this week

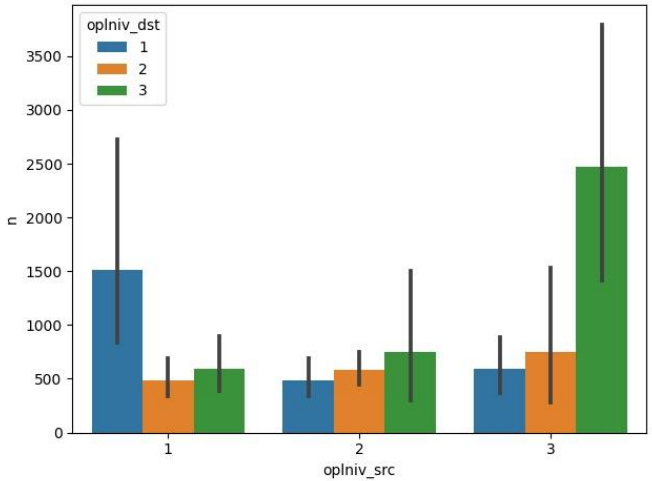
- More in depth Data investigation

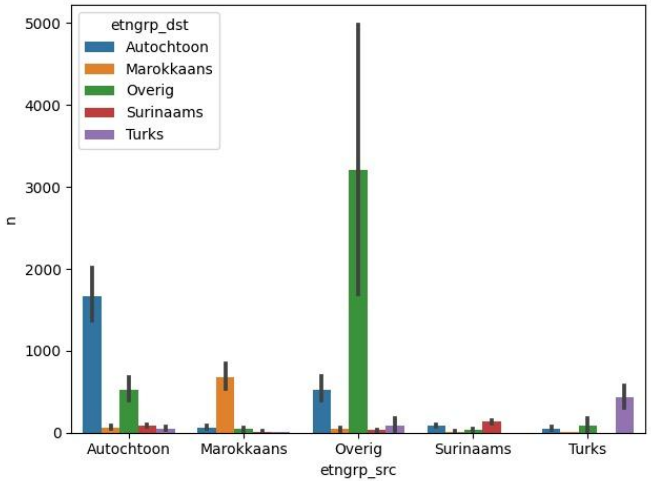
- Looked at the diversity of links per main group

- Read some literature on stochastic oriented actor model
  - Introduction to stochastic actor-based models for network dynamics (Snijders, van de Bunt, Steglichc, 2010)
  - Slides from a introduction to SAOM (Duke network analysis centre)

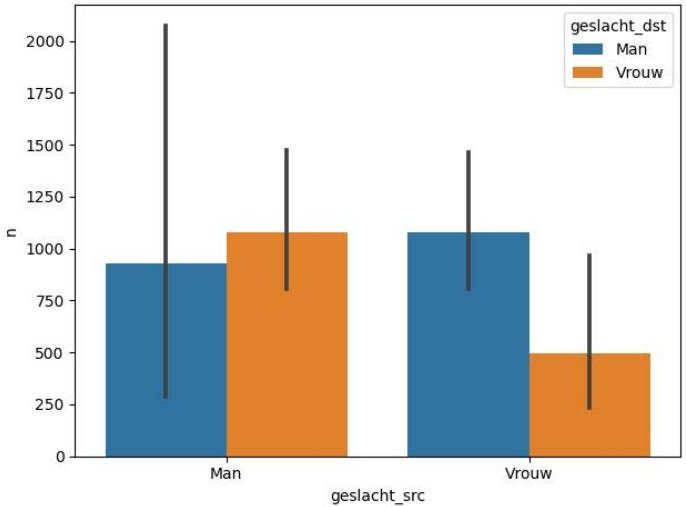
Household

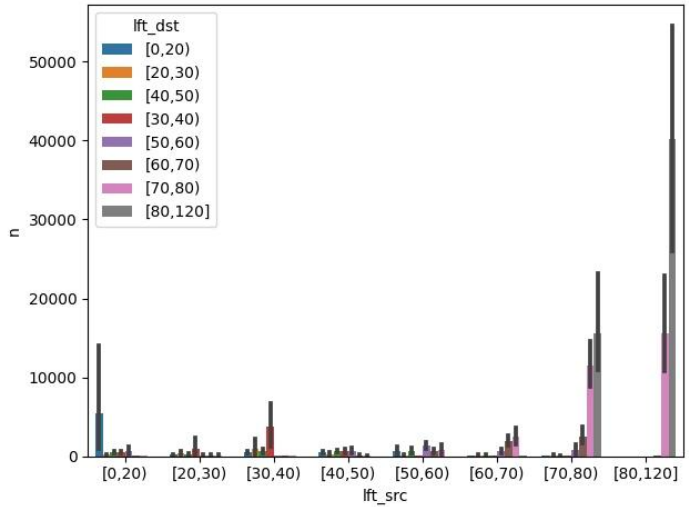
Normalize and statistical test look at education level



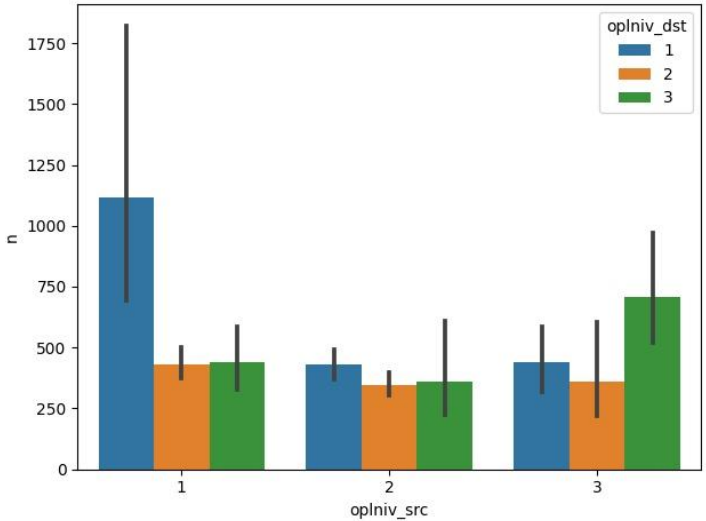


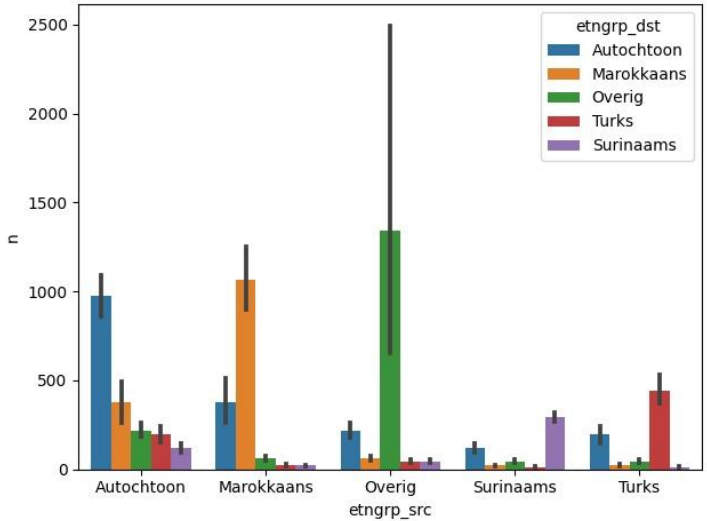
Household


geslacht_dst
Man
Vrouw
2000
1750
1500
1250
1000
750
500
250
0
n
Man
Vrouw
geslacht_src

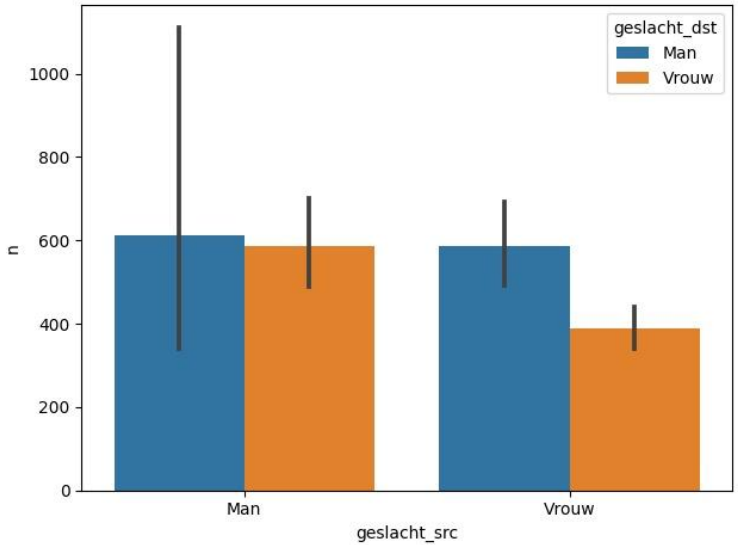

lft_dst
[0,20)
[20,30)
[40,50)
[30,40)
[50,60)
[60,70)
[70,80)
[80,120]
50000
40000
30000
20000
10000
0
n
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
[70,80)
[80,120]
lft_src

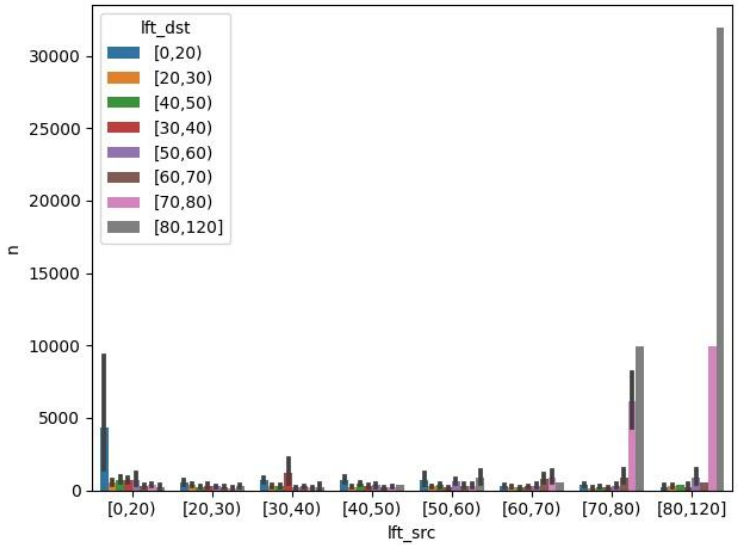
Family


oplniv_dst
1
2
3
n
1750
1500
1250
1000
750
500
250
0
1
2
3
oplniv_src

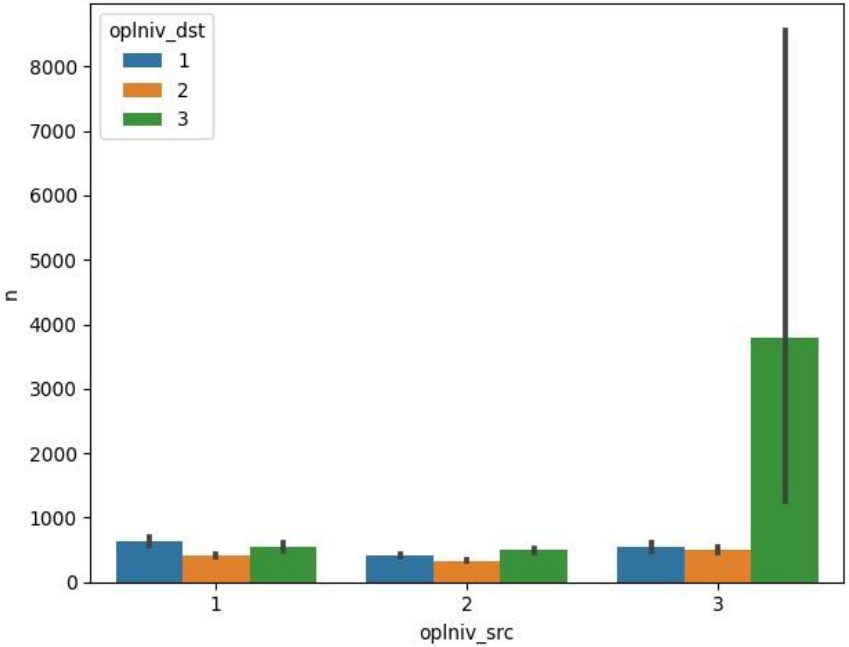

etngrp_dst
Autochtoon
Marokkaans
Overig
Turks
Surinaams
n
2500
2000
1500
1000
500
0
Autochtoon
Marokkaans
Overig
Surinaams
Turks
etngrp_src

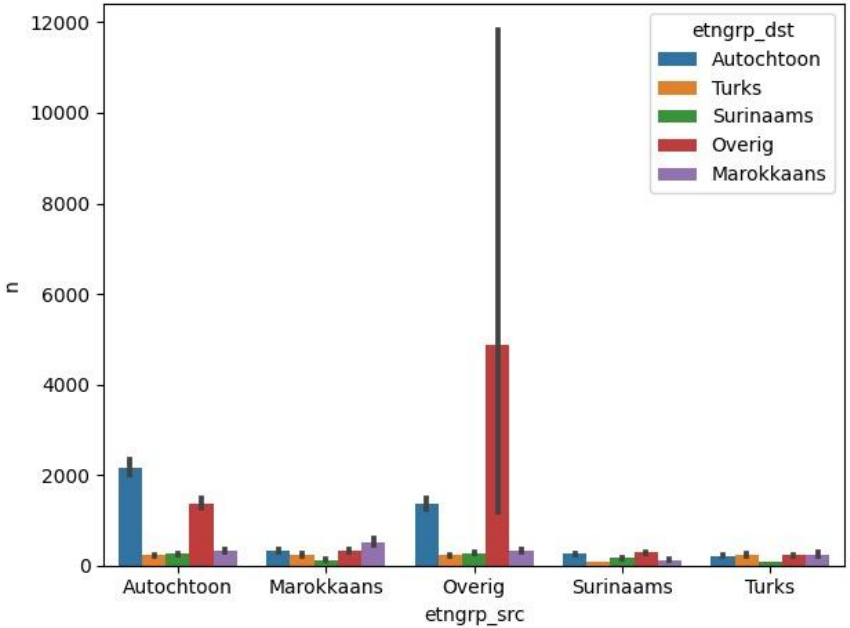
Family


geslacht_dst
Man
Vrouw
1000
800
600
400
200
0
n
Man
Vrouw
geslacht_src

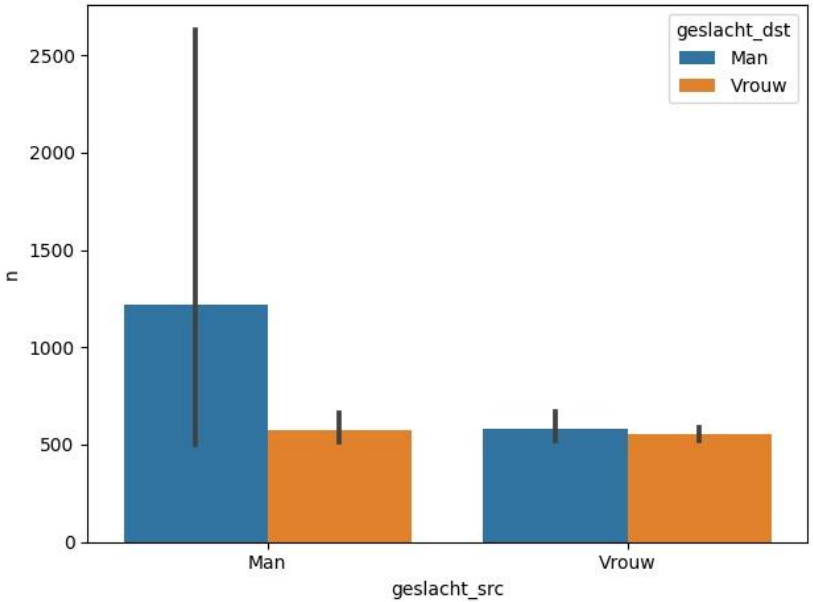

lft_dst
[0,20)
[20,30)
[40,50)
[30,40)
[50,60)
[60,70)
[70,80)
[80,120]
30000
25000
20000
15000
10000
5000
0
n
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
[70,80)
[80,120]
lft_src

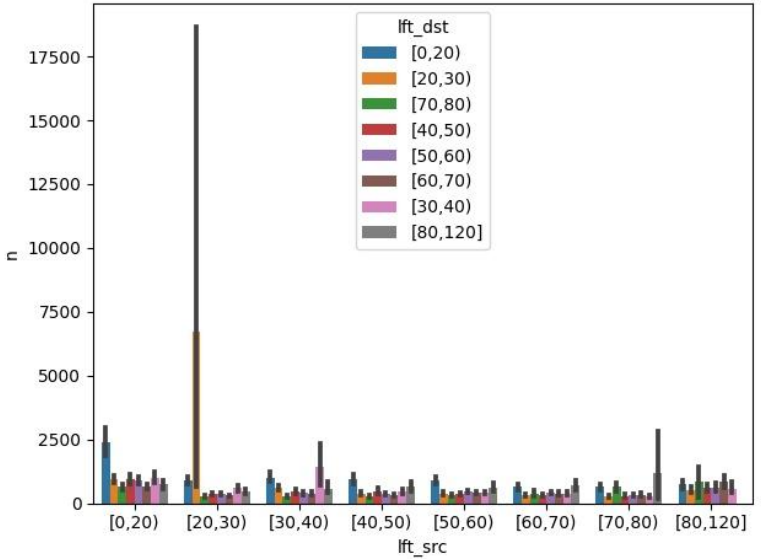
Neighbours


oplniv_dst
1
2
3
8000
7000
6000
5000
4000
3000
2000
1000
0
n
1
2
3
oplniv_src


etngrp_dst
Autochtoon
Turks
Surinaams
Overig
Marokkaans
12000
10000
8000
6000
4000
2000
0
n
Autochtoon
Marokkaans
Overig
Surinaams
Turks
etngrp_src

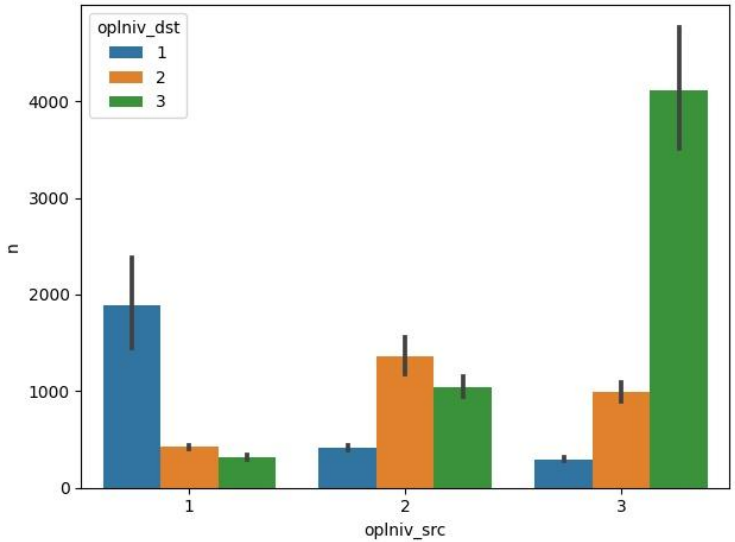
Neighbours

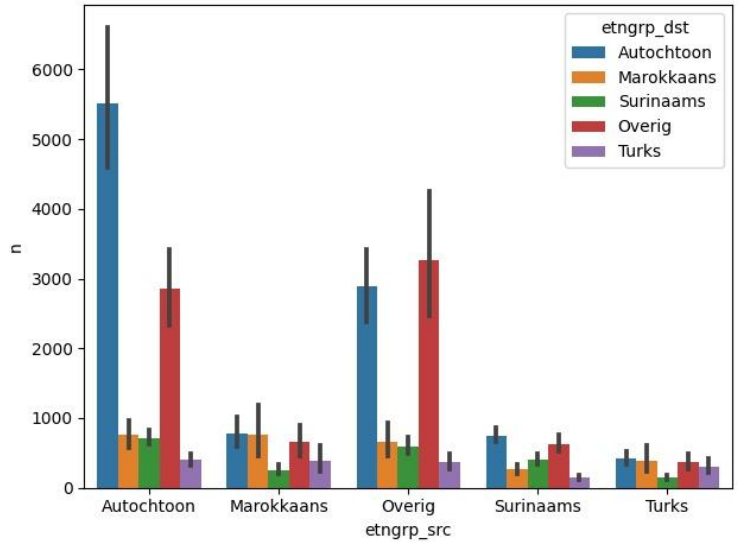

geslacht_dst
Man
Vrouw
2500
2000
1500
1000
500
0
n
Man
Vrouw
geslacht_src


lft_dst
[0,20)
[20,30)
[70,80)
[40,50)
[50,60)
[60,70)
[30,40)
[80,120]
17500
15000
12500
10000
7500
5000
2500
0
n
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
[70,80)
[80,120]
lft_src

Work/school

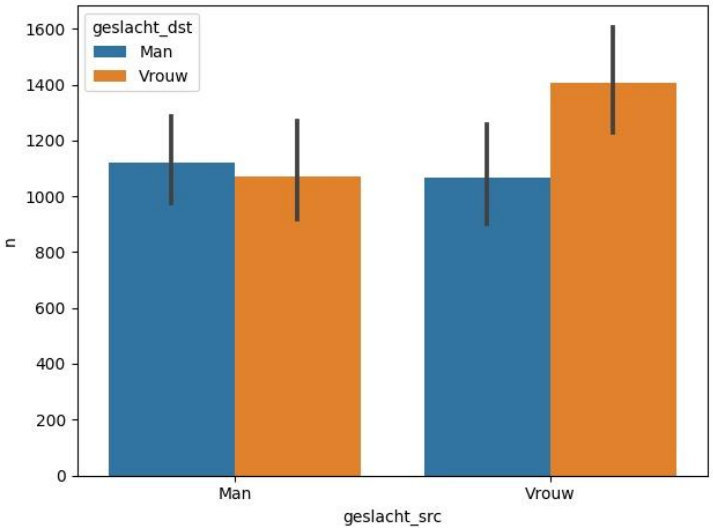
Heatmap visualisation

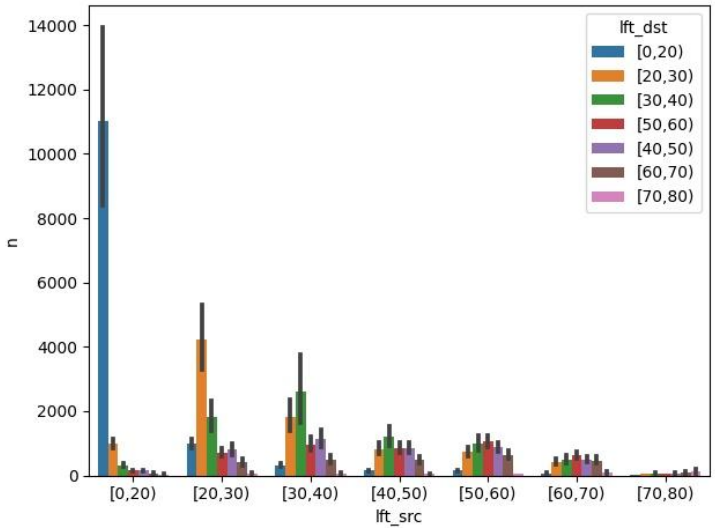




Work/school

entropy == > diversity





Stochastic oriented actor model

- Longitudinal study
- Initial network
- Multiple panel waves
- Make micro steps based on objective function
- Connections appear/disappear once at a time

- **Understand the measures**

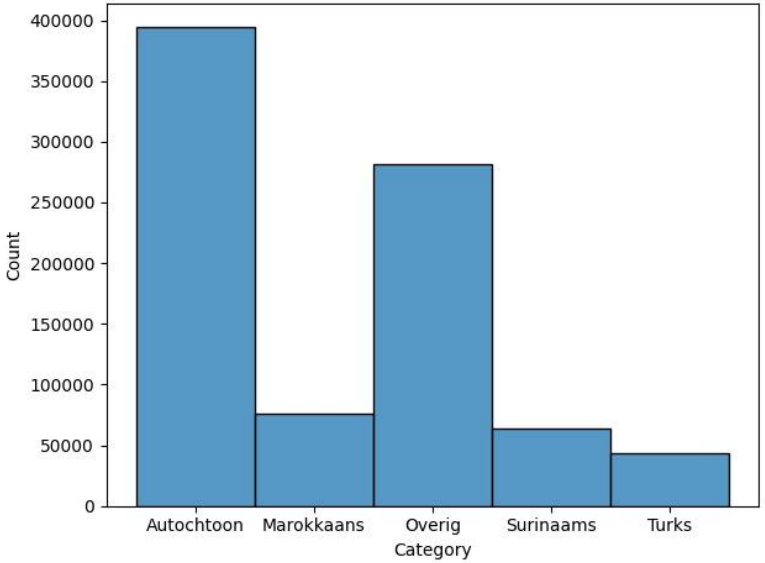
Planning to do

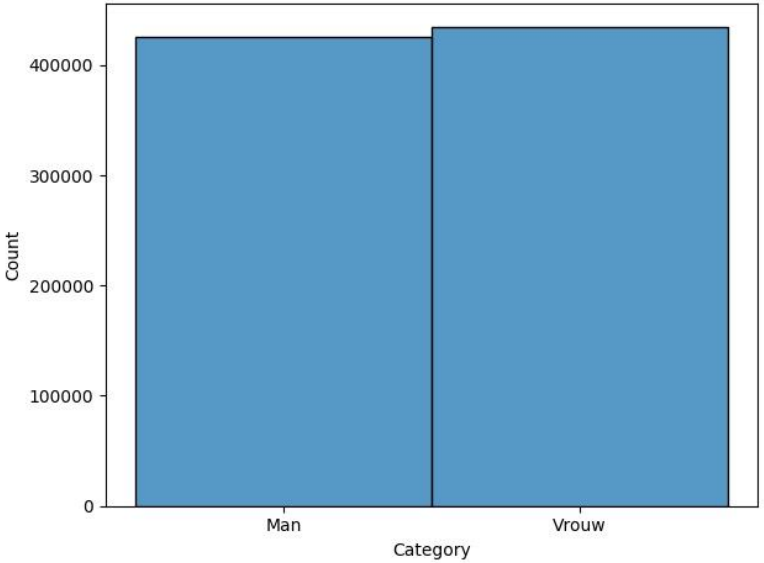
- Look more at stochastic actor oriented model (measurement)
- Think of possible implementation (Tom Snijder) ⇒
- Look at connections over time (by looking at the connections of each age category)
- City of Amsterdam distributions age etc.
- Measures of homophily
- R py2.

Done this week

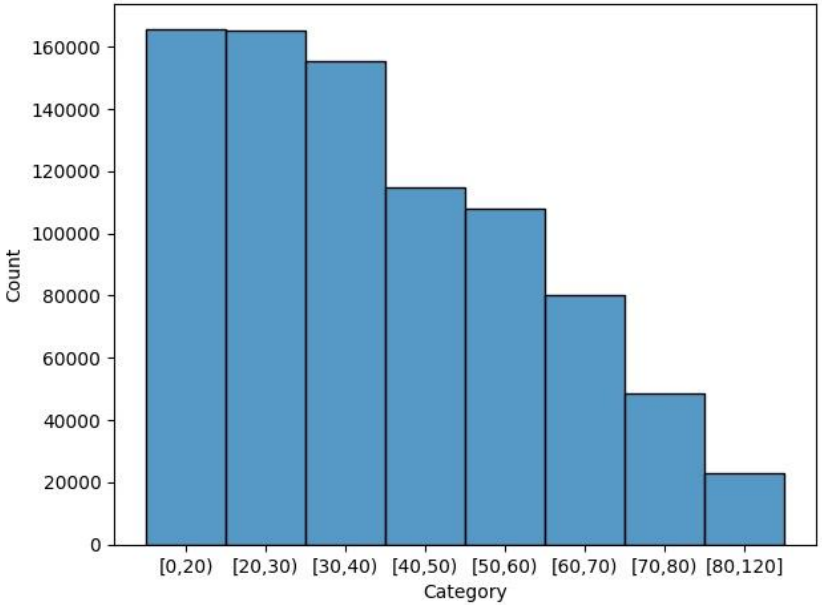
- Made distributions
-
- Looked at theils index
- Looked at literature about multi-layered networks

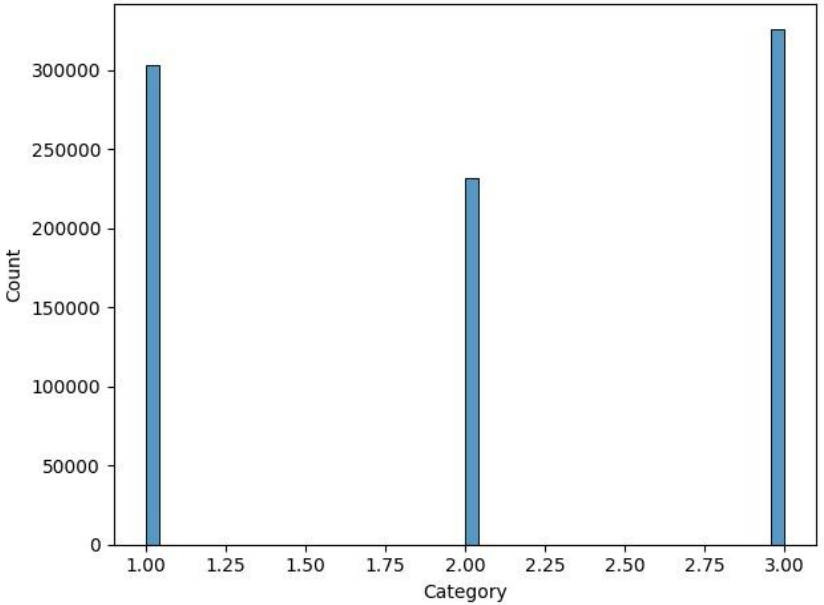
Distributions





Pile toghterg

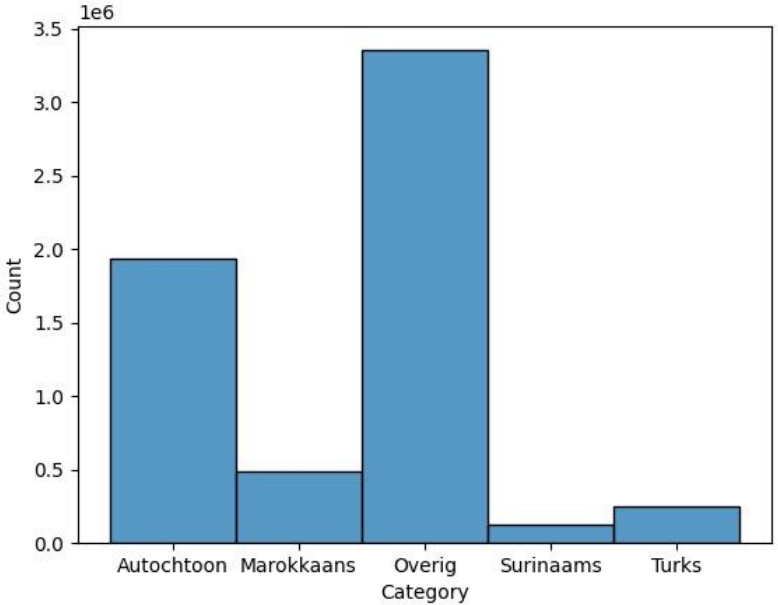


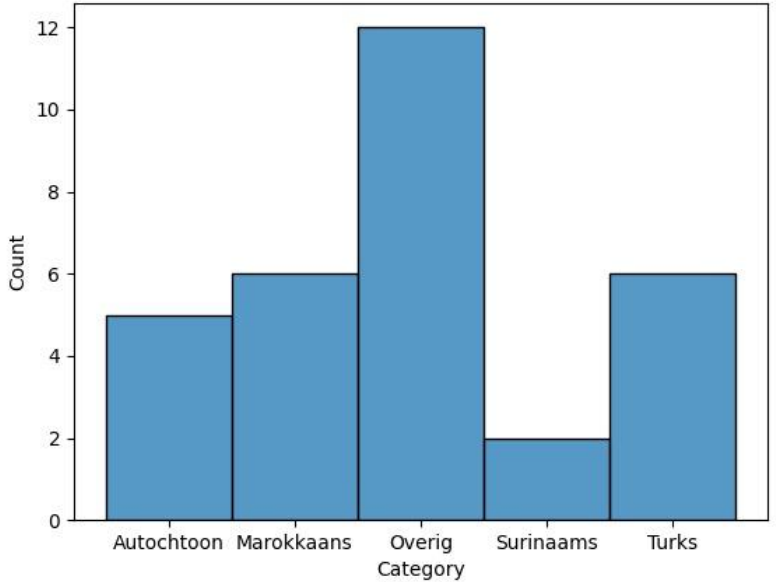


Connection distributions

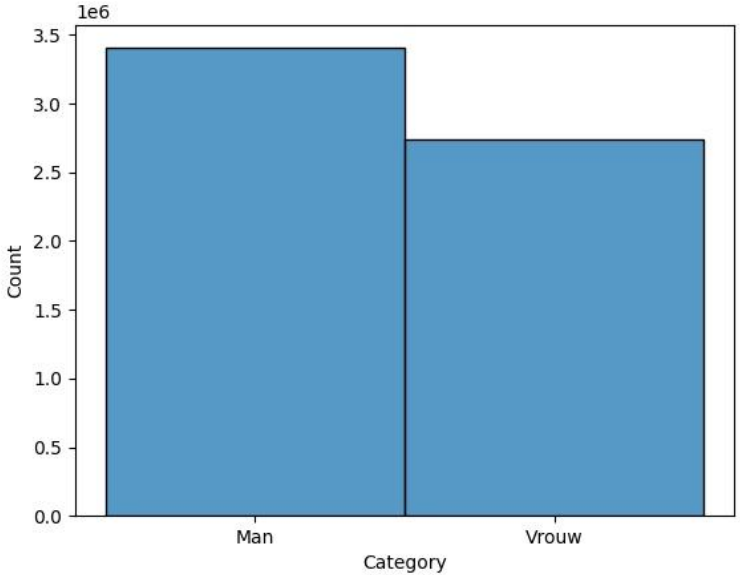
Total connections and connections per person

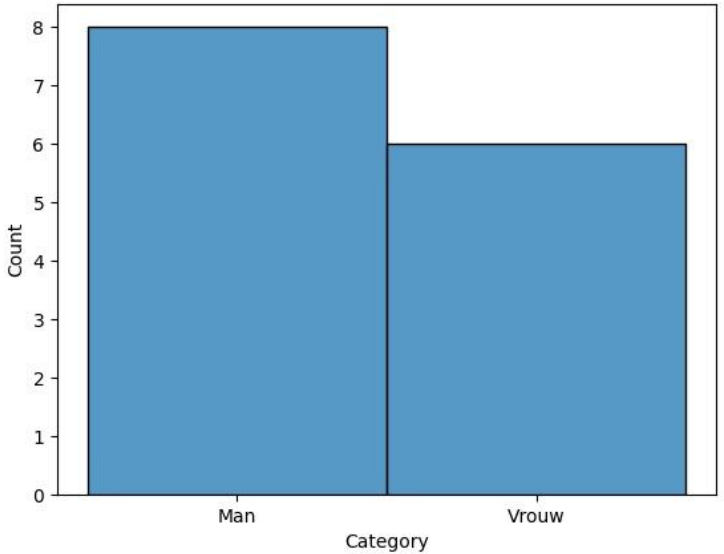
Household


1e6
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Count
Autochtoon
Marokkaans
Overig
Surinaams
Turks
Category

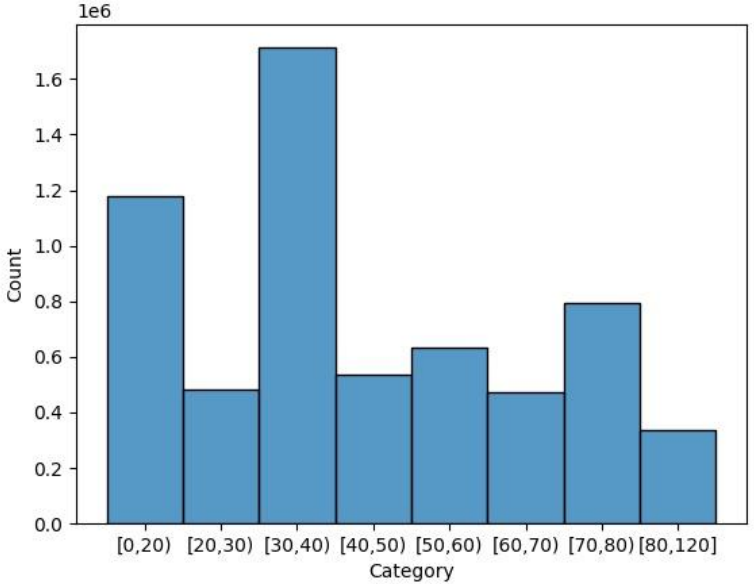

12
10
8
6
4
2
0
Count
Autochtoon
Marokkaans
Overig
Surinaams
Turks
Category

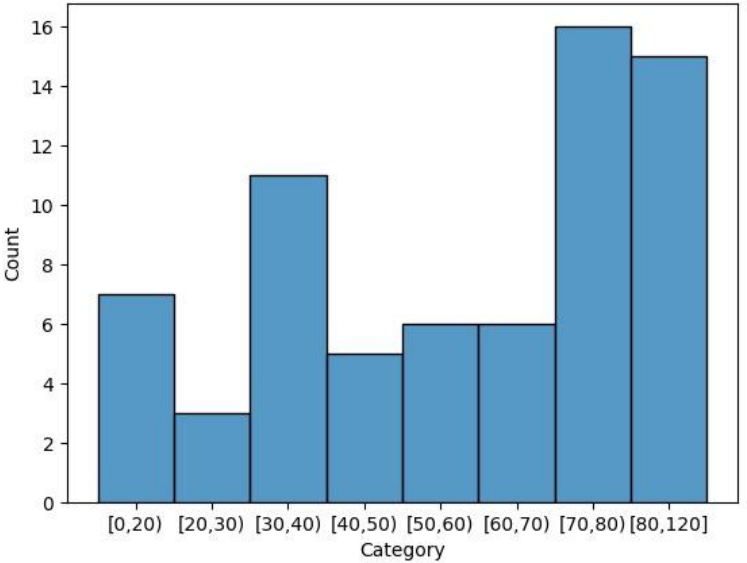
Household


1e6
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Count
Man
Vrouw
Category

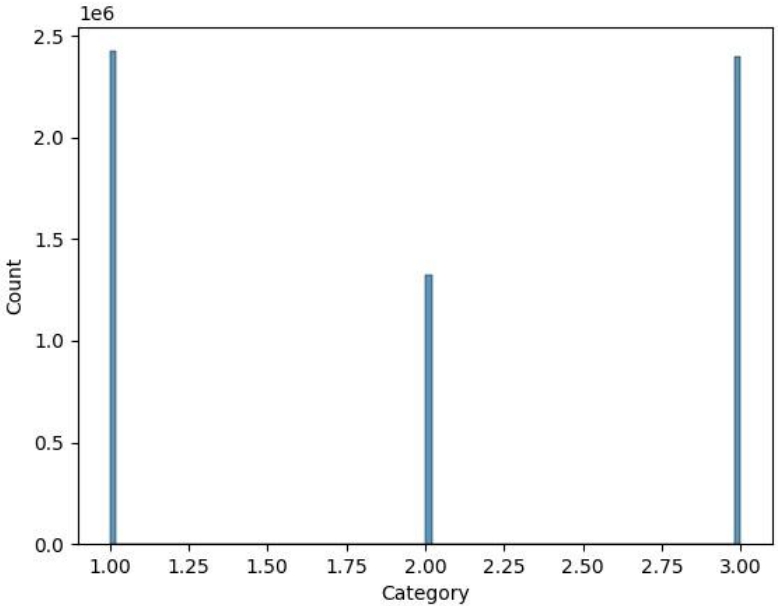

8
7
6
5
4
3
2
1
0
Count
Man
Vrouw
Category

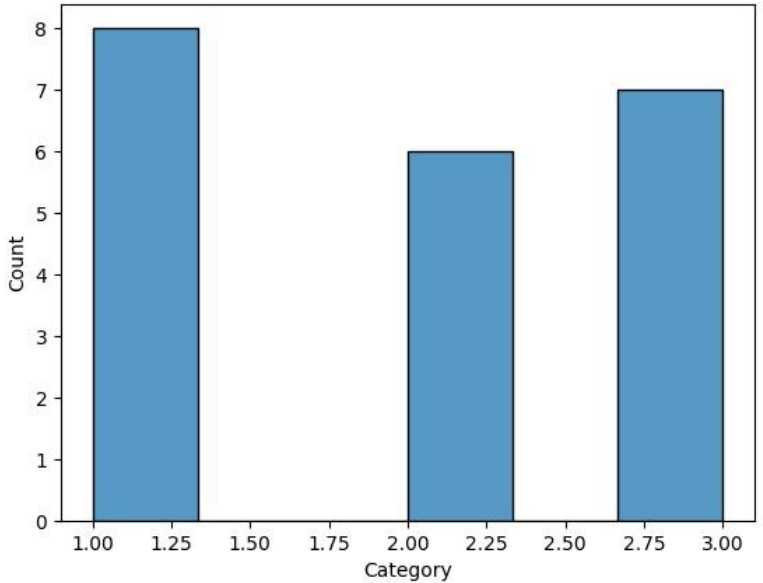
Household


1e6
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
[70,80)
[80,120]
Category

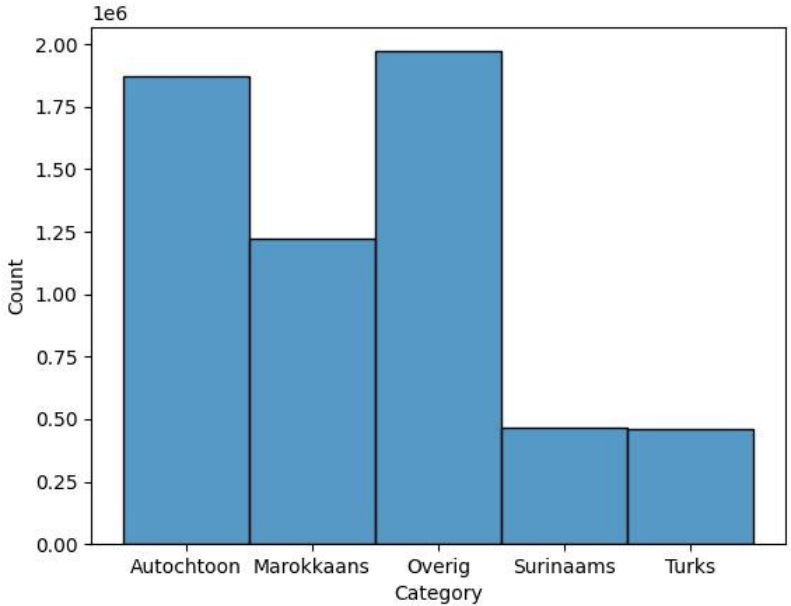

16
14
12
10
8
6
4
2
0
Count
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
[70,80)
[80,120]
Category

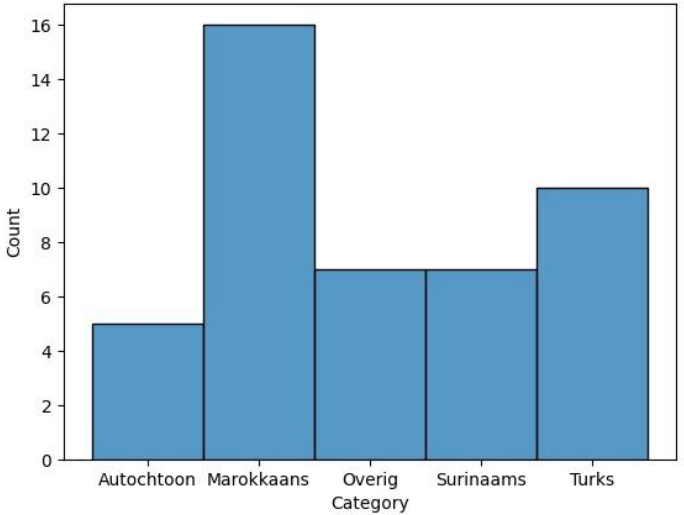
Household


1e6
2.5
2.0
1.5
1.0
0.5
0.0
Count
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
Category

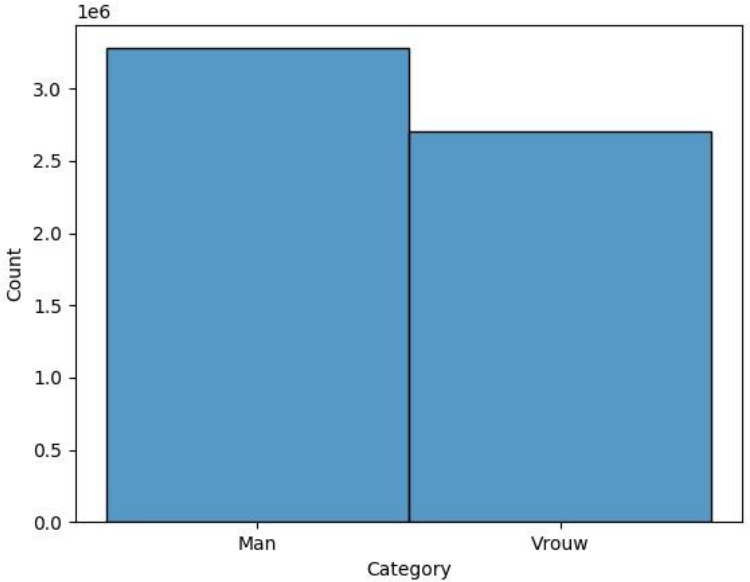

8
7
6
5
4
3
2
1
0
Count
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
Category

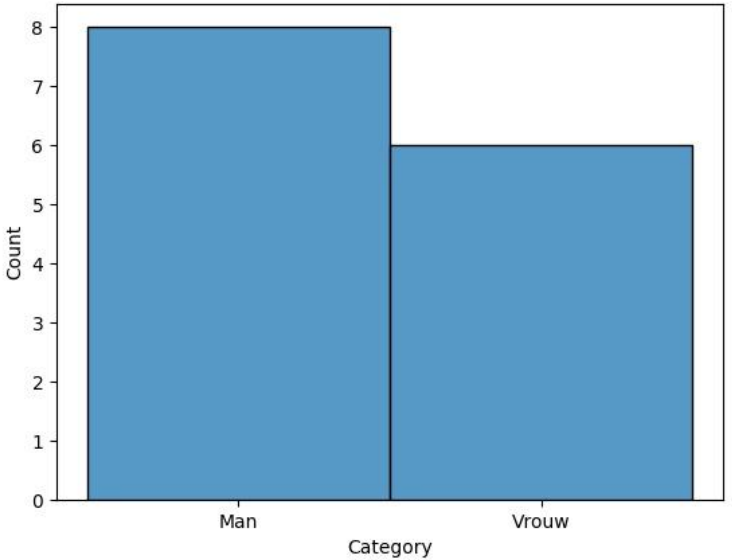
Familie


1e6
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Count
Autochtoon
Marokkaans
Overig
Surinaams
Turks
Category

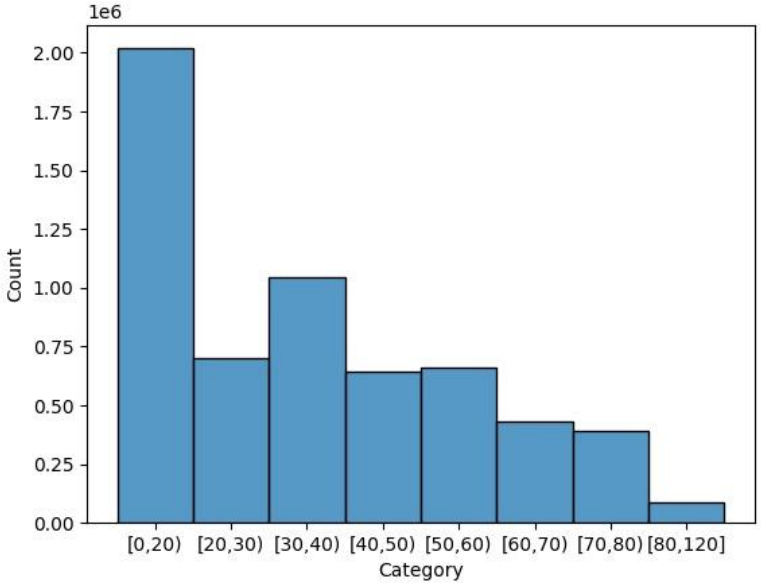

16
14
12
10
8
6
4
2
0
Count
Autochtoon
Marokkaans
Overig
Surinaams
Turks
Category

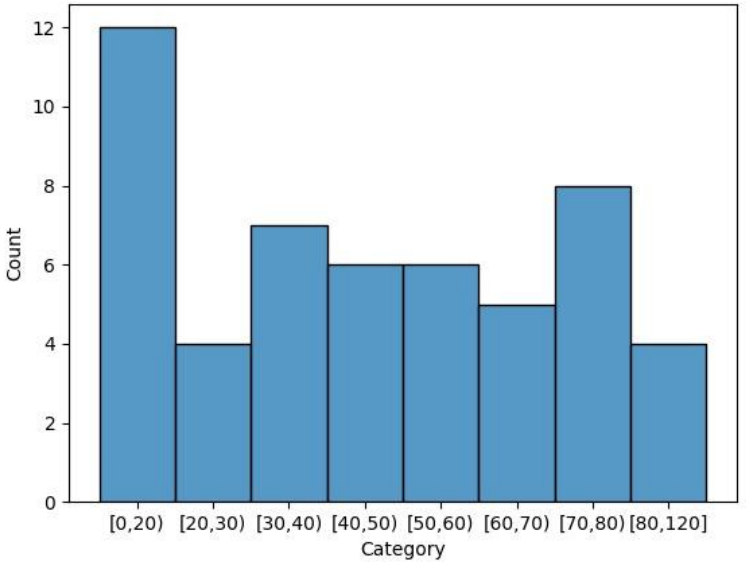
Familie


1e6
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Count
Man
Vrouw
Category

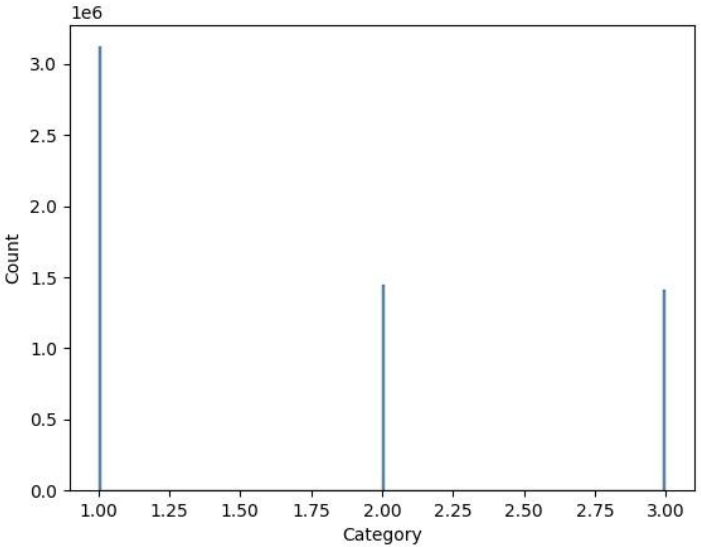

8
7
6
5
4
3
2
1
0
Count
Man
Vrouw
Category

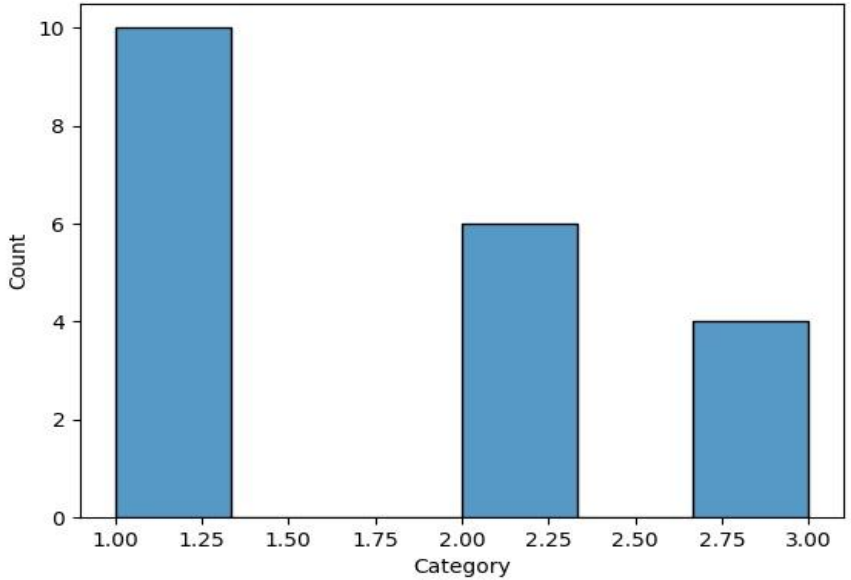
Familie


1e6
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Count
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
[70,80)
[80,120]
Category

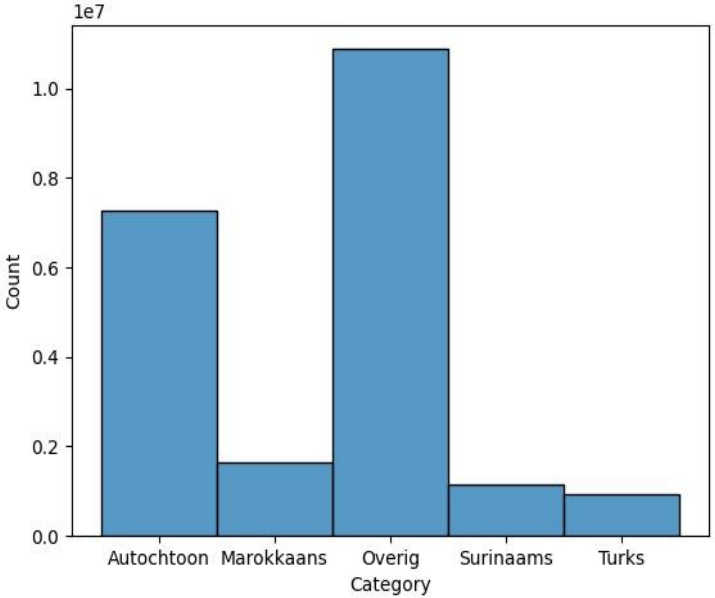

12
10
8
6
4
2
0
Count
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
[70,80)
[80,120]
Category

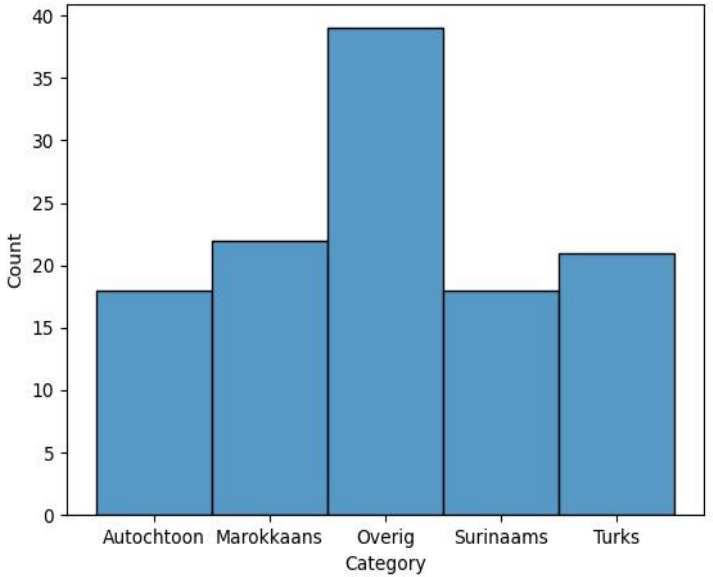
Familie


1e6
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Count
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
Category

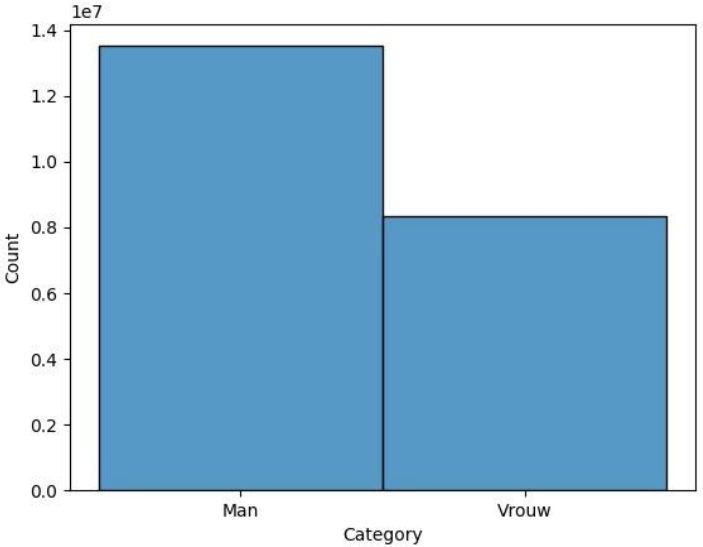

10
8
6
4
2
0
Count
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
Category

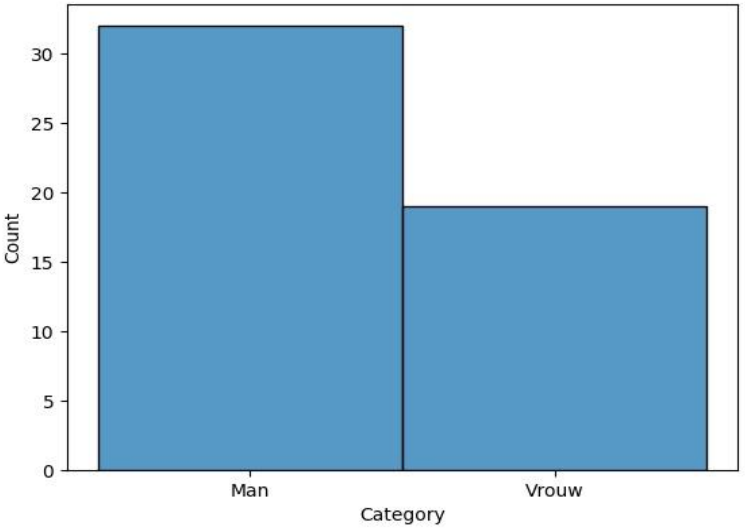
Neighbours


1e7
1.0
0.8
0.6
0.4
0.2
0.0
Count
Autochtoon
Marokkaans
Overig
Surinaams
Turks
Category

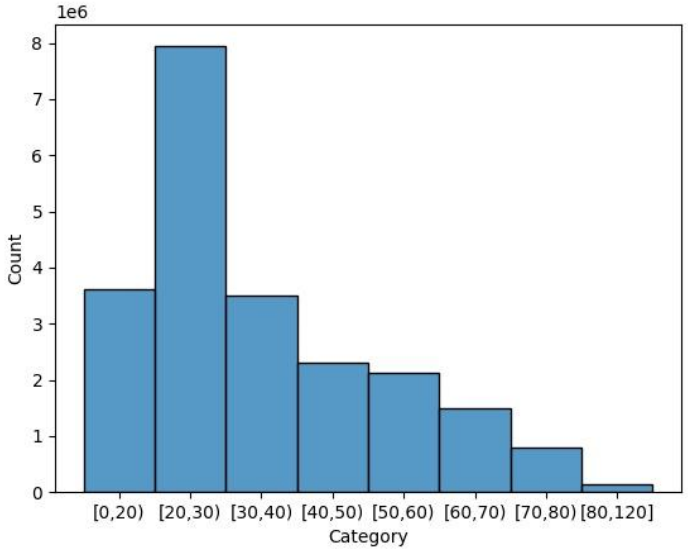

40
35
30
25
20
15
10
5
0
Count
Autochtoon
Marokkaans
Overig
Surinaams
Turks
Category

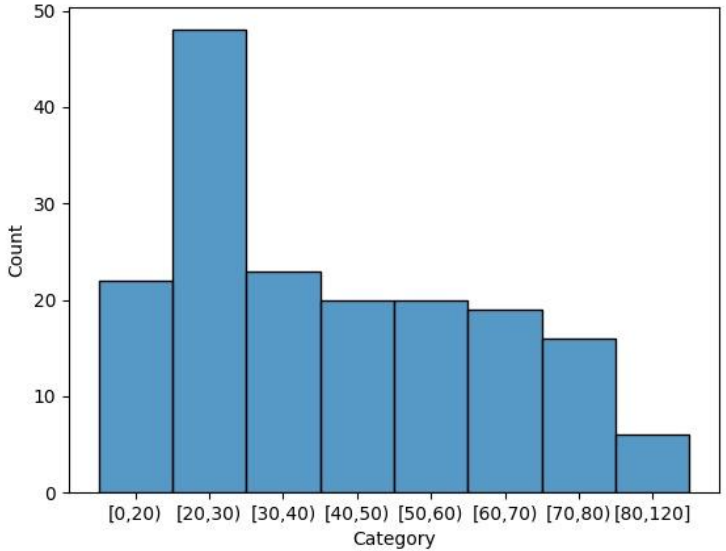
Neighbours


1e7
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
Man
Vrouw
Category

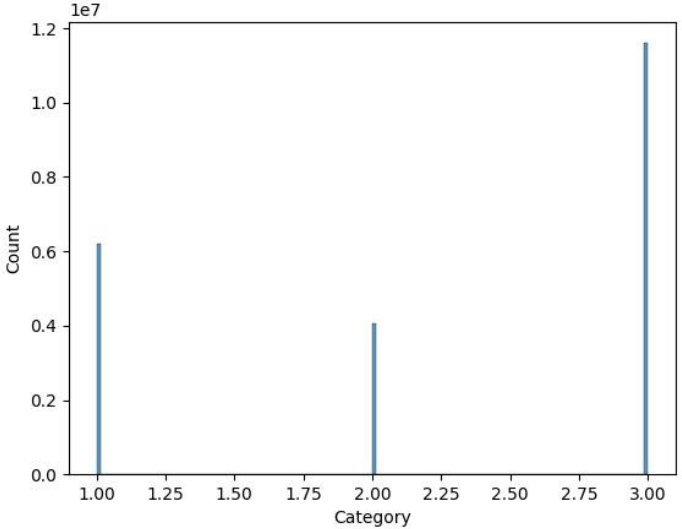

30
25
20
15
10
5
0
Count
Man
Vrouw
Category

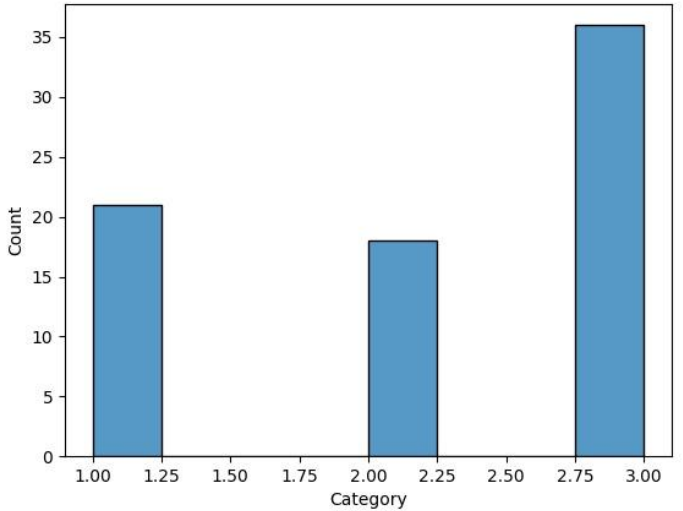
Neighbours


1e6
8
7
6
5
4
3
2
1
0
Count
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
[70,80)
[80,120]
Category

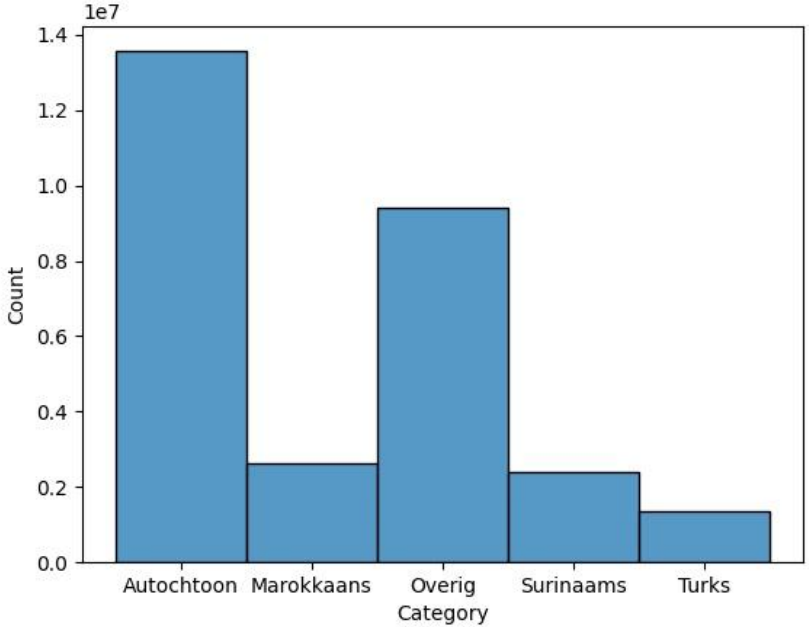

50
40
30
20
10
0
Count
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
[70,80)
[80,120]
Category

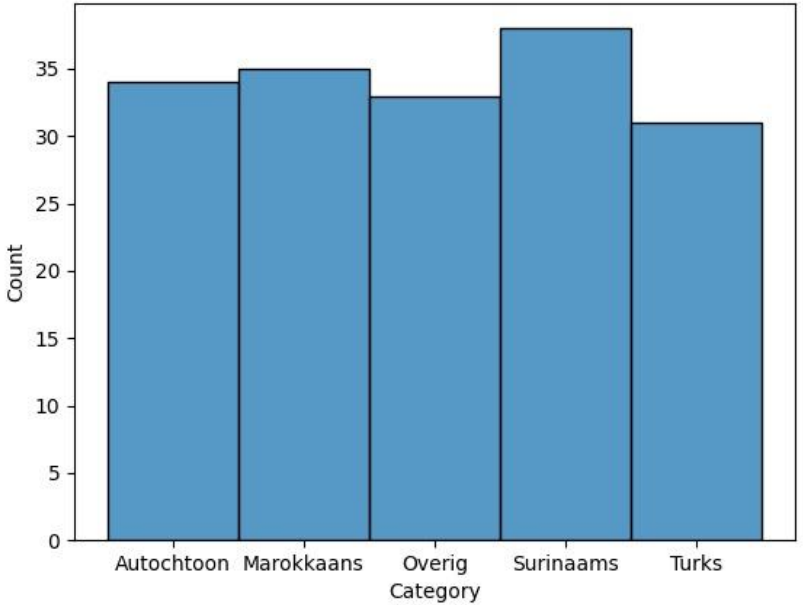
Neighbours


1e7
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
Category

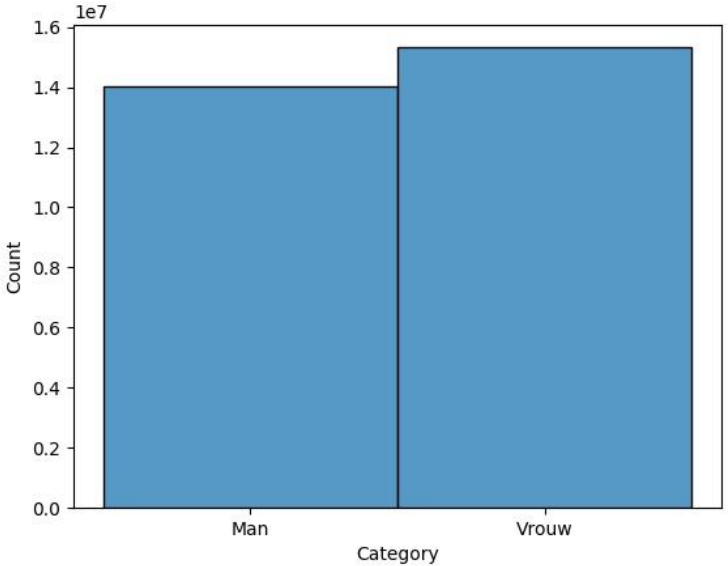

35
30
25
20
15
10
5
0
Count
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
Category

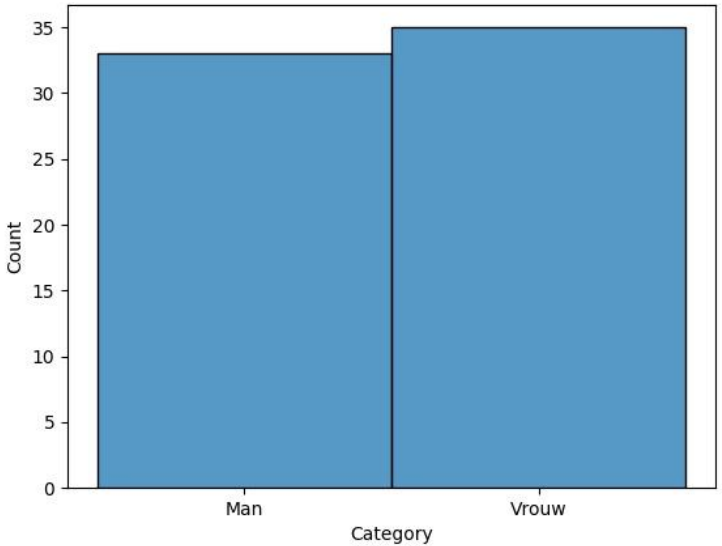
Work/School


1e7
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
Autochtoon
Marokkaans
Overig
Surinaams
Turks
Category

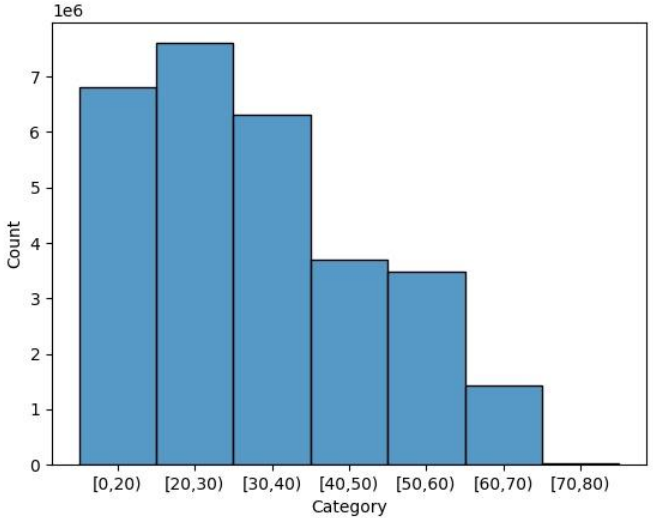

35
30
25
20
15
10
5
0
Count
Autochtoon
Marokkaans
Overig
Surinaams
Turks
Category

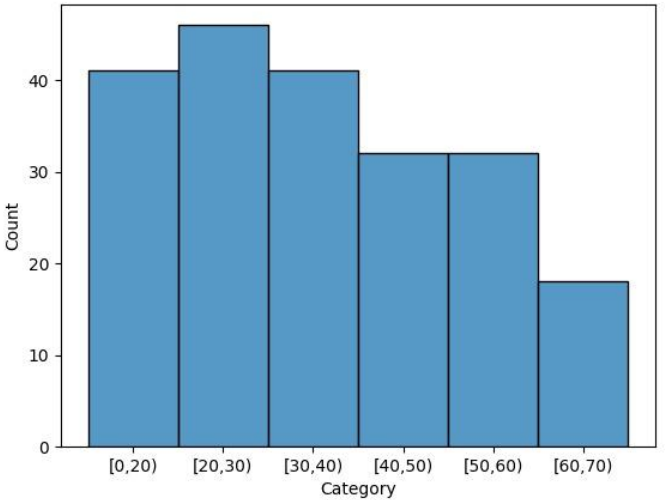
Work/School


1e7
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
Man
Vrouw
Category

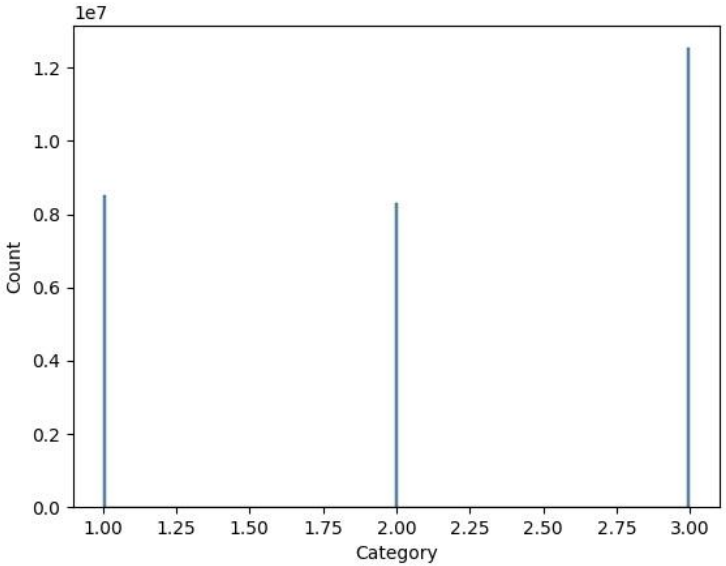

35
30
25
20
15
10
5
0
Count
Man
Vrouw
Category

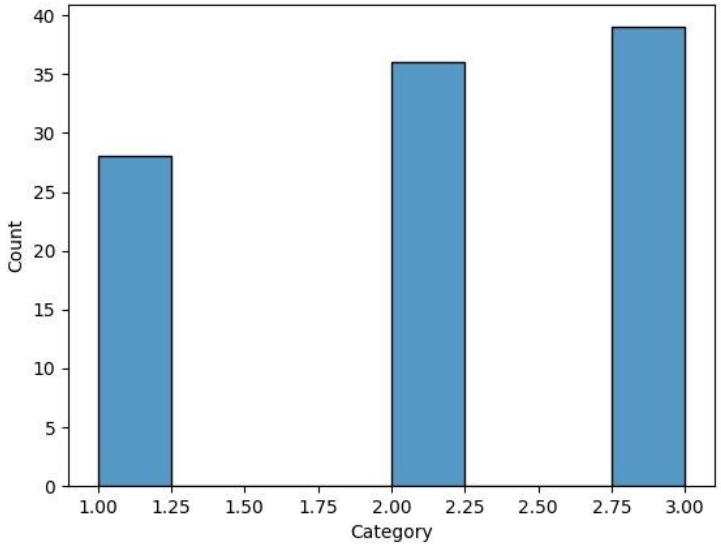
Work/School


1e6
7
6
5
4
3
2
1
0
Count
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
[70,80)
Category


40
30
20
10
0
Count
[0,20)
[20,30)
[30,40)
[40,50)
[50,60)
[60,70)
Category

Work/School


1e7
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Count
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
Category


40
35
30
25
20
15
10
5
0
Count
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
Category

Theils index

Multi layer network

- Multilayer Networks Structure and Function (Ginestra Bianconi)

- Towards real-world complexity: an introduction to multiplex networks (Kyu-Min Lee, Byungjoon Min, and K.-I. Goh)

Email Loes

Segregation literature (amsterdam)

Know what overig represents

Institutional connections

Overlap household age and etnicity

Mutual informations → Models

https://nowosad.github.io/post/raceland-bp1/

https://www.nature.com/articles/srep44981

loss function

Complete diversity

Discussed before christmas break

- Making network connections based on links instead of probabilities
- When decreasing the nodes with x, decrease edges with (x**2/2)
- Making a popularity parameter

What I did in the last few days

- Reading more literature about multilayer social networks
- Mostly from the book of Dickens et al.,
- Which give a inside in how a multilayered network differs from a monolayered network and what the different kind of analysis one can make

Building the network

- Made all the layers of the multilayer network but than a smaller fraction
- Nodes/10, edges/(10**2/2)
- The neighbour layer and Workschool layer are random while household and family are random but symmetric

- 1 + links node has / total_links take it to a power p to make popularity parameter

Some analysis on the monolayer level

	edges	nodes	connected_nodes	avg indegree	max in degree	avg out degree	max out degree	avg clustercoefficient	max cluster coefficient	reciprocity
huishouden	141572	86100	50459	2.805684	57_1534, [0,20), Overig, 1	2.805684	57_1534, [0,20), Overig, 1	0.000191	44_664, [50,60), Overig, 3	1.000000
familie	136988	86100	56482	2.425339	57_2460, [0,20), Overig, 1	2.425339	57_2460, [0,20), Overig, 1	0.000251	128_132, [50,60), Autochtoon, 3	0.999898
buren	437425	86100	85508	5.115603	68_614, [20,30), Overig, 3	5.115603	68_484, [20,30), Overig, 3	0.000736	127_207, [50,60), Autochtoon, 2	0.009099
werkschool	587055	86100	79422	7.391592	52_133, [0,20), Overig, 2	7.391592	58_99, [0,20), Overig, 2	0.000219	34_221, [60,70), Overig, 2	0.000371

Looking for a way to make the monolayers into a multilayered network

- Found a package which does that but it had some problems with installing
- Otherwise maybe trying to implement it myself

Other thoughts

- Family and household are symmetric but neighbour and work/school should give a higher probability of connecting when there is already one connection
- When X -> Y then 0.8 probability that Y -> X
- Not sure how much a scale free network makes sense as the data is not scale free (10 neighbours, 100 closest colleague/classmate) but might still give interesting results

- Spatial data?
- look at the parameters and objective function
- look at the degree distribution (scale-free) for the bins
- Look at degree distibution whole network
-

Previously discussed

- Constructing network based on edges between two groups instead of probability
- Making popularity parameter based on Barabasi

What I did this week

- Reading survey literature (Book of Dickson and book of Bianconi)
- Implemented popularity parameter
- Optimizing code so it runs faster

Small analysis on total network

	edges	nodes	connected_nodes	avg indegree	max in degree	avg out degree	max out degree	reciprocity
huishouden	283790	86100	66134	4.291136	57	4.291136	57	1.000000
familie	274634	86100	72231	3.802163	57	3.802163	57	1.000000
buren	875310	86100	85903	10.189516	68	10.189516	68	0.017464
werkschool	1174412	86100	80119	14.658346	58	14.658346	58	0.000715

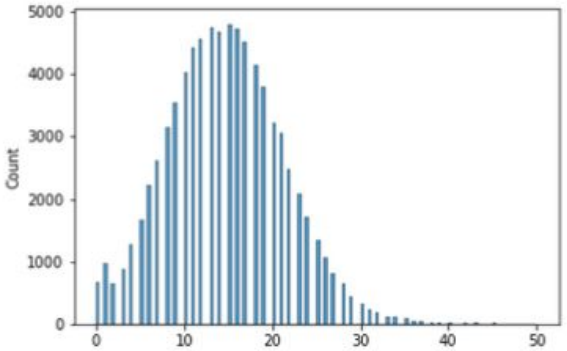
Not really representative

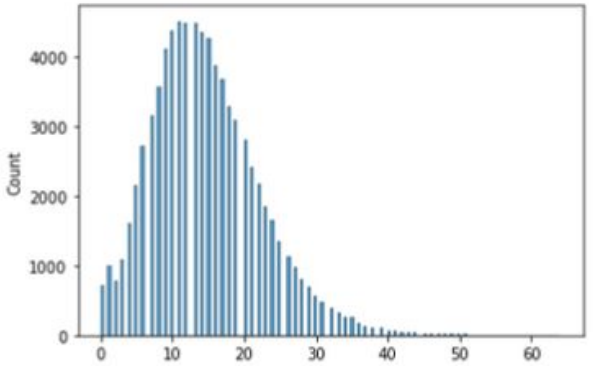
(got different values with whole network which corresponds to the data)

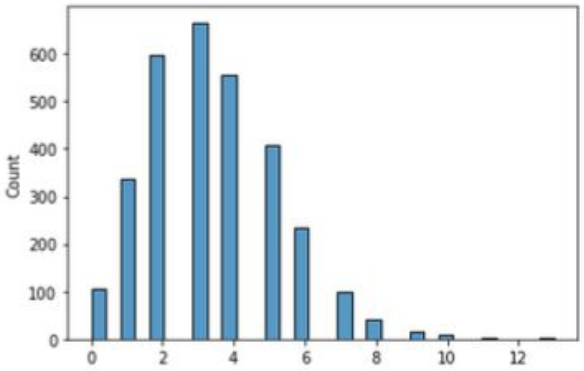
Implementation of popularity parameter

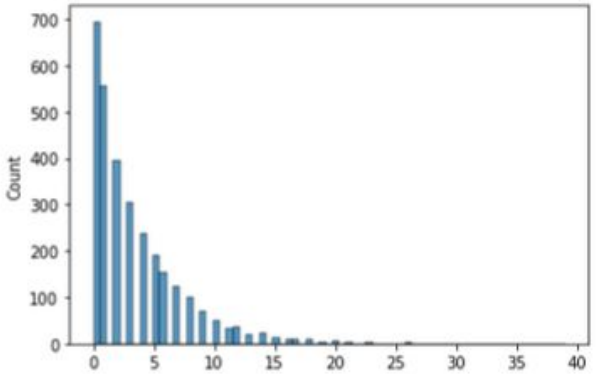
- Implemented a weight
- Each time a node gets chosen the weight gets increased by 1
- Also made a power function out of it so we can choose from a popularity parameter 0 to 1

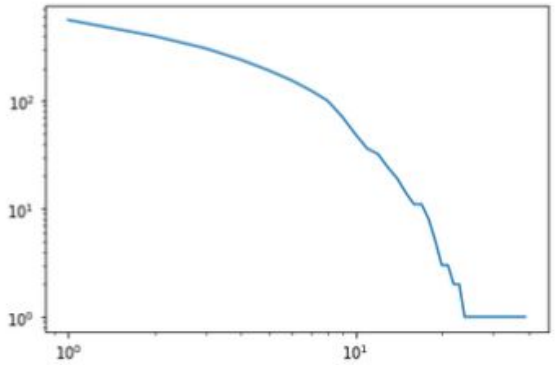
Distributions










10^2
10^1
10^0
10^0
10^1

Optimizing code

- I hashed all strings
- Used as much low level packages as possible (such as numpy)
- Discarded pandas so I do not have to iterate over these rows
- Discarded classes

- I can now run the whole network in a minute, but only without popularity parameter

Problem

- Distribution between groups looks like a exponential distribution and not a power distribution. Can maybe be fixed by increasing the popularity (increase probability of being chosen)

- I can now run the full network but not with popularity parameter (makes the programme slow)

To do

- Try to improve popularity parameter so the code can run full network
- Looking at a reciprocity parameter (now I have full reciprocity for family and household but maybe give a certain probability instead)
- Looking at spatial data and making the network based on that
- Reading more literature (finishing both books can be helpful)
- Make mono layers into a multiplex/multilayer network
- Starting already a little bit on writing the thesis so I keep structure
- Dunbar number

X -> Y then Y -> X 100%

Mainly focused on Barabasi algorithm

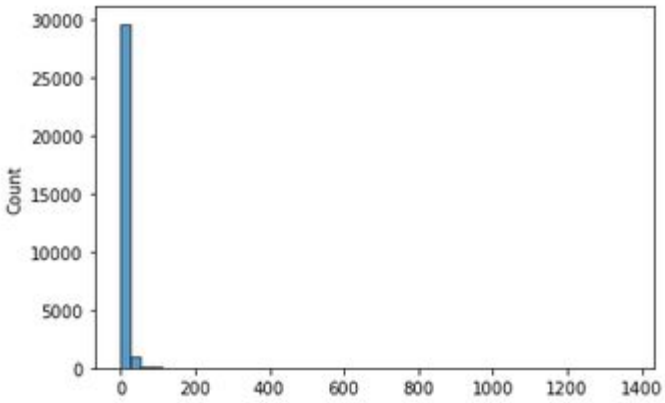
- Looking at the Barabasi algorithm
  - Making it faster
  - Making it work so it is scale-free and not exponential
  - Looking at the exponential distribution - normal distribution behaviour
  - Dunbar number

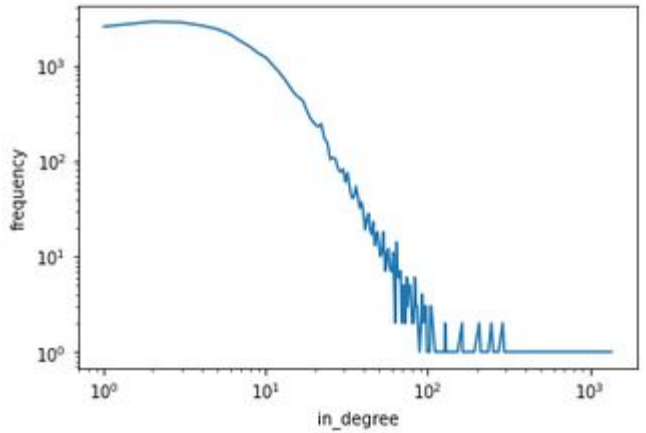
Scale-free network

- The way I implemented the Barabasi algorithm does not show the scale-free property
- So I changed it to the following:
  - Start by choosing a random destination node from the destination group
  - Put this node inside a destination bin
  - Take a random node from the source group
  - Connect random node from source group with a random node from the destination bin (which has in the beginning only 1 node inside)
  - Add the chosen node from the bin again in the Bin (so we increase its probability of being chosen)
  - Add a random other node from the destination group to the destination bin

Checking if sale-free property is met

- Example between the same group ('[0,20), Native, Man, 1')
- Looks scale free

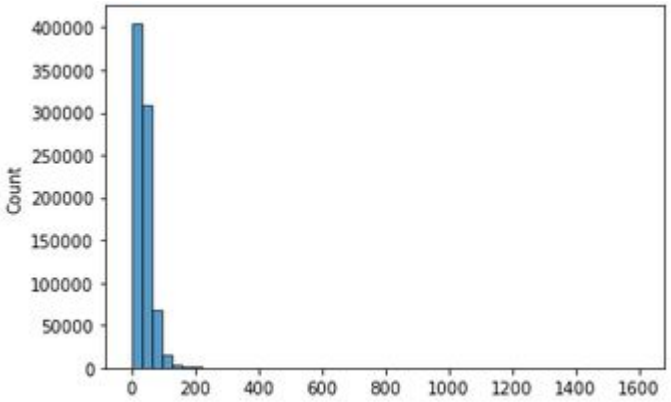


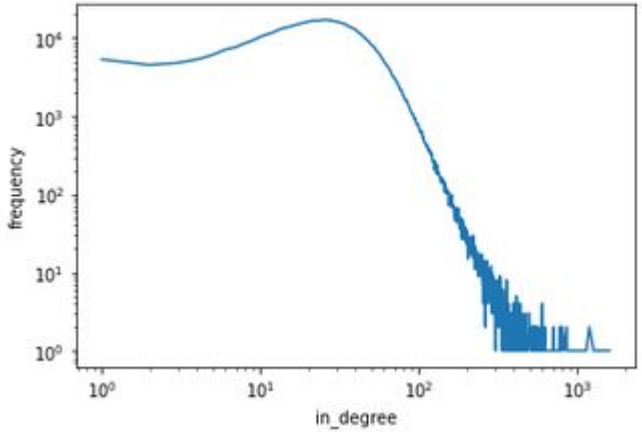


Checking if sale-free property is met

- Using the package power_law to make sure
  - Jeff Alstott, Ed Bullmore, Dietmar Plenz. (2014). powerlaw: a Python package for analysis of heavy-tailed distributions
- Comparing fit between power law and exponential
  - power law significant
- Power law - truncated power law
  - No significant difference
- Power law/truncated power law - lognormal
  - No significant difference
  - But not big problem as Alslott et al., show that even well known powerlaw data fit both
  - Also the paper of Broido and Clauset show that lognormal are even more common and almost all powerlaws also fit lognormal
    - Broido, A.D., Clauset, A. Scale-free networks are rare. *Nat Commun* **10,** 1017 (2019).

Looking at the whole network

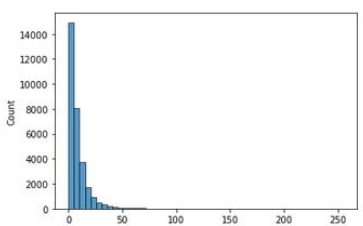


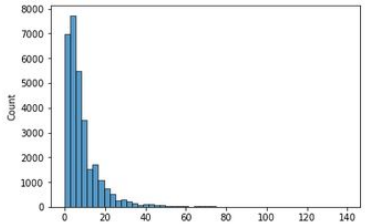


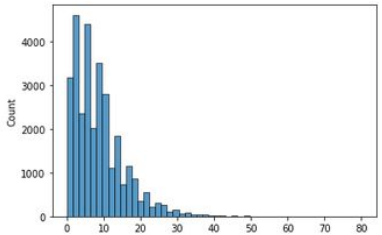
- Statistics are the same as with the ('[0,20), Native, Man, 1') group

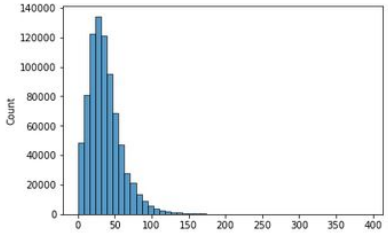
Exponential distribution

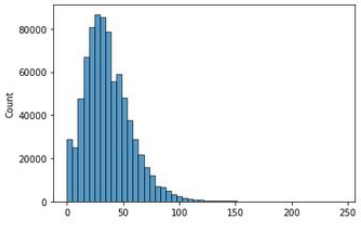
- When we increase the amount of initial nodes in the destination bin, we see that the distribution between groups start following an exponential distribution and the total network will look more like a normal distribution
- E. of workschool data on groups ('[0,20), Native, Man, 1') with 1,5 and 50%

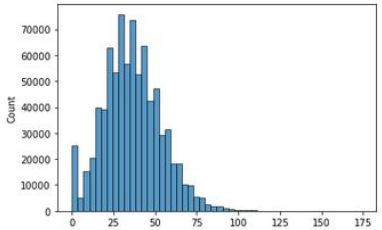












Speeding up the network initialization

- Because I changed the weighted random.choice function with a normal random.choice function the process speeded up a lot.
- Both Barabasi and random network can now be runned on the local computer in 3-4 minutes

Dunbar number in the scale-free network (work/school)

- It has been proposed to lie between 100 and 250, with a commonly used value of 150.
- With a destination bin of 1% only 66 persons have a higher in degree value than 250, the largest has a in degree of 394 making it more plausible 1% initial nodes more plausible than only 1 node which results in
- highest in-degree of 1987 links and 963 nodes above 250 in-degree

- Interesting is that with random network the maximum in_degrees lays on 152

To do

Sharing

- I will upload my code to github so everyone can see it
- Also I will put all the presentations together so we have a overview

Initial network

- Still looking at reciprocity
- Making multilayered network
- Looking at spatial data
- Age distribution

Dynamical network

- Also thinking about possible dynamics
- Maybe brainstorm on the possibilities

Notes

- Instead one put all the nodes based on the random parameter
- High percentile to dunbar
- Algorithms that produce scale-free properties when you have two groups
  - Check if there is a paper already on this topic

- Only Age the population in static network
  - age distribution in Amsterdam in group
  - Age population into the future ⇒ remove nodes when dead
  - Change links when getting older or stay the same
  - **Show network dynamics move one edge to the other**
    - **Check how new age group compares to old age group**

Discussed before christmas break

- Making network connections based on links instead of probabilities
- When decreasing the nodes with x, decrease edges with (x**2/2)
- Making a popularity parameter

What I did in the last few days

- Reading more literature about multilayer social networks
- Mostly from the book of Dickens et al.,
- Which give a inside in how a multilayered network differs from a monolayered network and what the different kind of analysis one can make

Building the network

- Made all the layers of the multilayer network but than a smaller fraction
- Nodes/10, edges/(10**2/2)
- The neighbour layer and Workschool layer are random while household and family are random but symmetric

- 1 + links node has / total_links take it to a power p to make popularity parameter

Some analysis on the monolayer level

	edges	nodes	connected_nodes	avg indegree	max in degree	avg out degree	max out degree	avg clustercoefficient	max cluster coefficient	reciprocity
huishouden	141572	86100	50459	2.805684	57_1534, [0,20), Overig, 1	2.805684	57_1534, [0,20), Overig, 1	0.000191	44_664, [50,60), Overig, 3	1.000000
familie	136988	86100	56482	2.425339	57_2460, [0,20), Overig, 1	2.425339	57_2460, [0,20), Overig, 1	0.000251	128_132, [50,60), Autochtoon, 3	0.999898
buren	437425	86100	85508	5.115603	68_614, [20,30), Overig, 3	5.115603	68_484, [20,30), Overig, 3	0.000736	127_207, [50,60), Autochtoon, 2	0.009099
werkschool	587055	86100	79422	7.391592	52_133, [0,20), Overig, 2	7.391592	58_99, [0,20), Overig, 2	0.000219	34_221, [60,70), Overig, 2	0.000371

Looking for a way to make the monolayers into a multilayered network

- Found a package which does that but it had some problems with installing
- Otherwise maybe trying to implement it myself

Other thoughts

- Family and household are symmetric but neighbour and work/school should give a higher probability of connecting when there is already one connection
- When X -> Y then 0.8 probability that Y -> X
- Not sure how much a scale free network makes sense as the data is not scale free (10 neighbours, 100 closest colleague/classmate) but might still give interesting results

- Spatial data?
- look at the parameters and objective function
- look at the degree distribution (scale-free) for the bins
- Look at degree distibution whole network