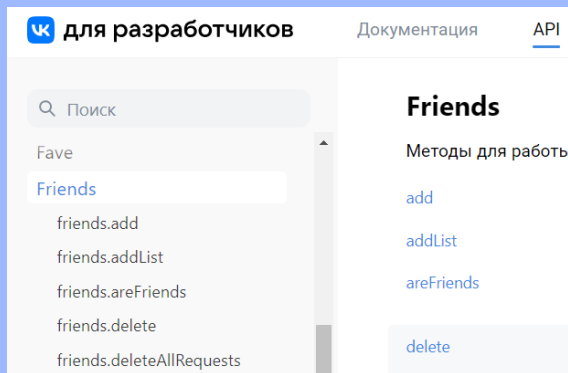


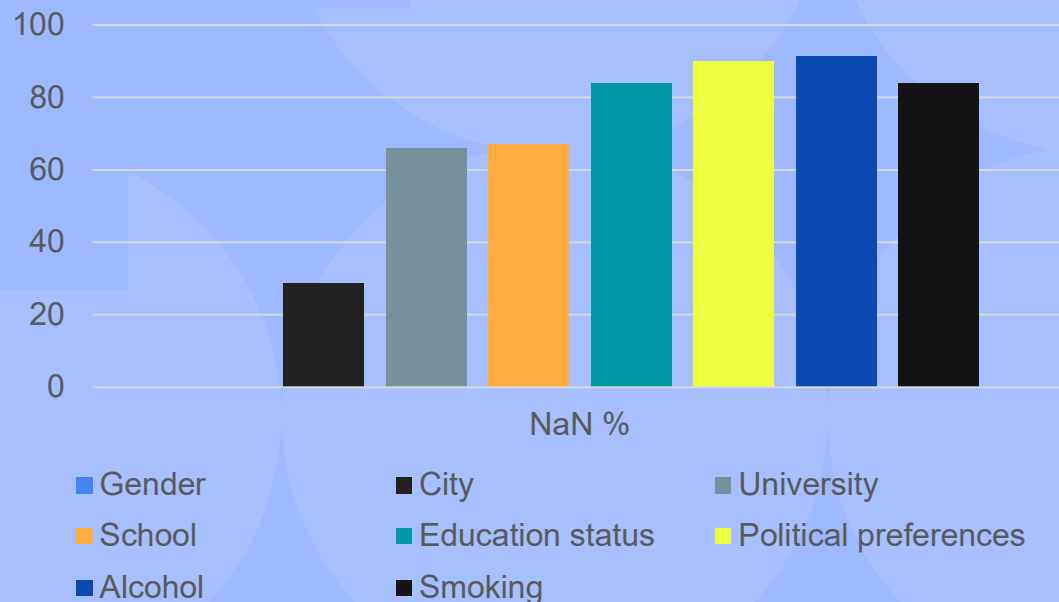
# Social Network Analysis Project

# Data collection



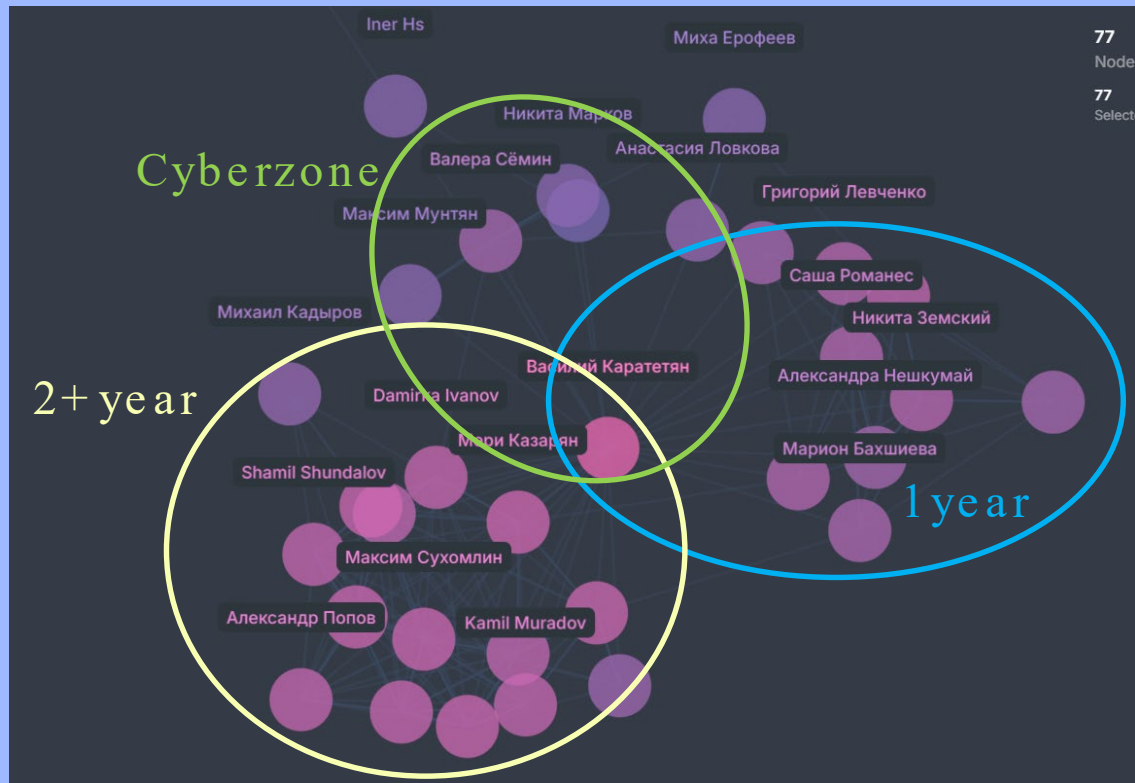
VK API

## Missing values



- ✦ Total friends: 98 (4 are inactive)
- ✦ Show on graph: 77 (the rest don't have connections)
- ✦ Largest connected component size - 73
- ✦ The graph is undirected, homogeneous, unweighted

# Summary

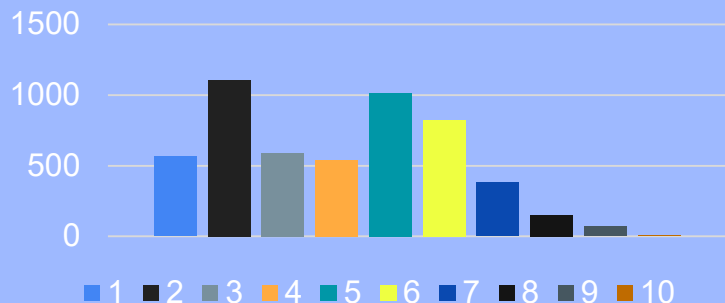


## The largest connected components statistics

- ✦ Nodes – 73
- ✦ Edges – 283
- ✦ Diameter – 10
- ✦ Radius – 5

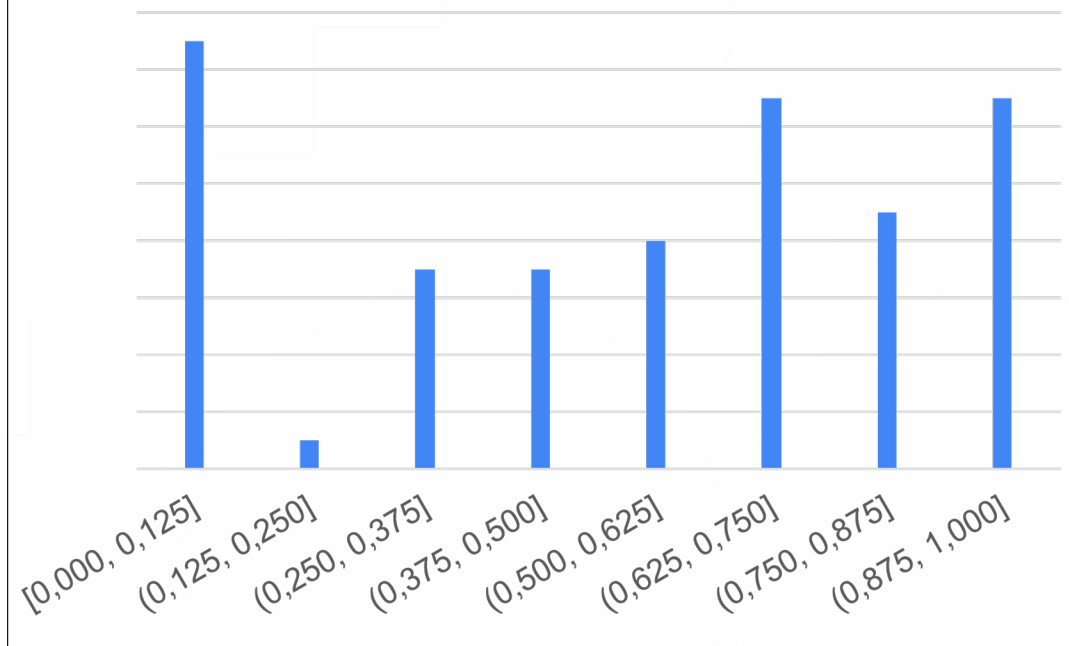
# Summary

## Distribution of shortest path lengths



Average shortest path length:  
4.06

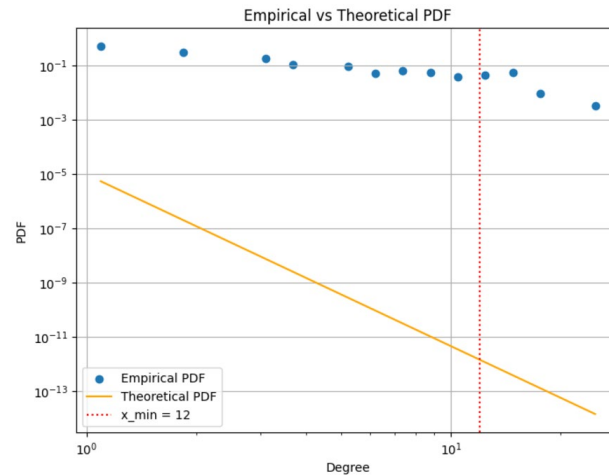
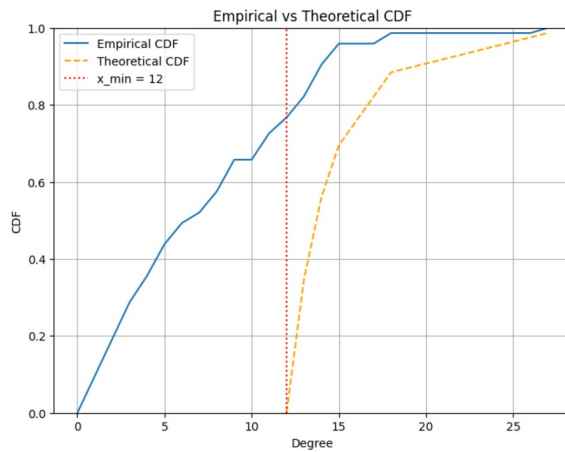
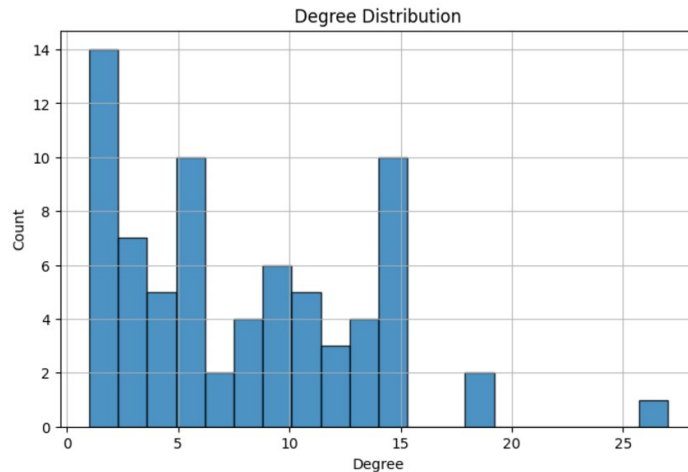
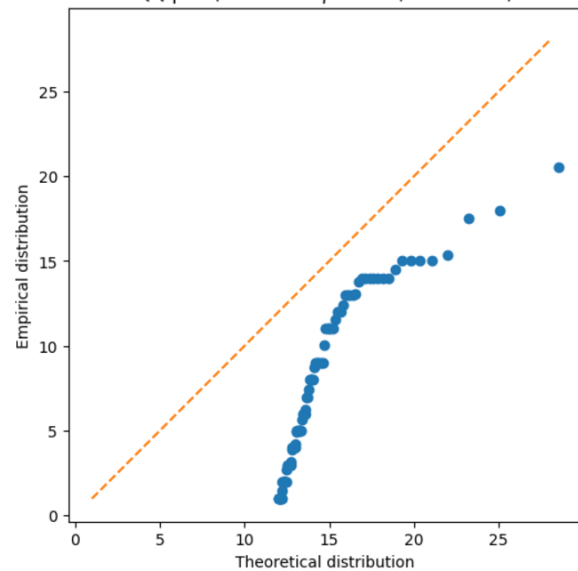
## Distribution of local clustering coefficients



Transitivity: 0.63  
Average local clustering coefficient:  
0.537

# Power law

QQ plot (estimated  $\gamma = 6.32$ ,  $x_{min} = 12$ )



# Random graphs

$N$  – number of nodes,  $K$   
 – number of edges (in the original graph)

Erdős-Rényi:  $n = N, p = \frac{2K}{n(n-1)}$

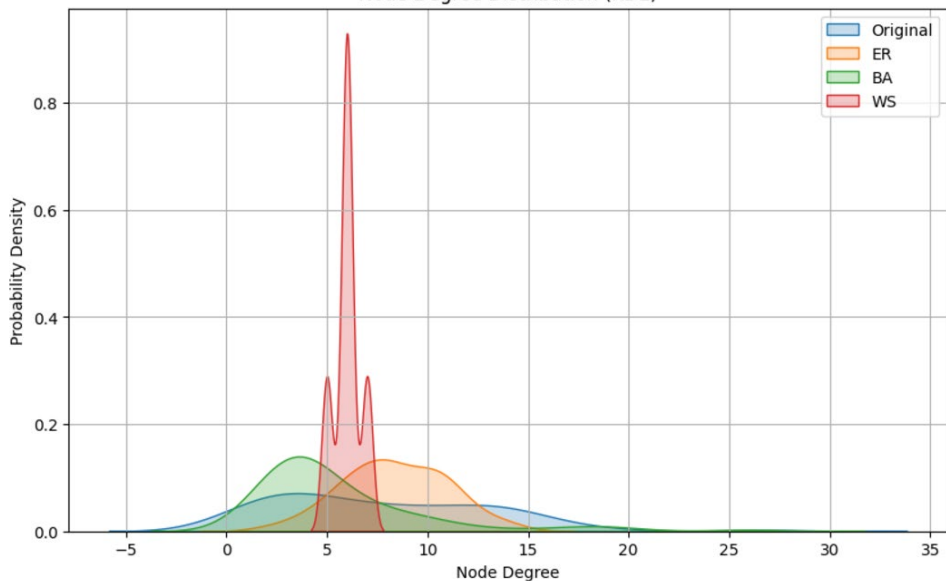
Barabasi-Albert:  $n = N, m = K \text{ div } N$

Watts-Strogatz:  $n = N, k = \frac{2K}{N}, p_{\text{rewire}} = 0.1$

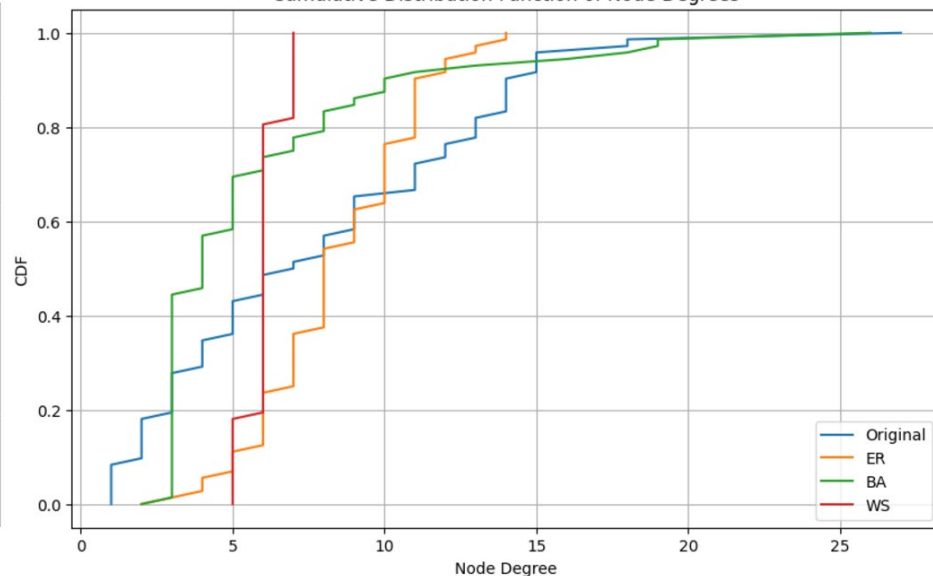
Metric	Original	ER	BA	WS
Radius	5	3	3	5
Diameter	10	4	4	8
Transitivity	0.63	0.11	0.13	0.49
Avg local clustering coefficient	0.54	0.10	0.17	0.53
Avg shortest path length	4.06	2.3	2.49	3.86

# Random graphs

Node Degree Distribution (KDE)



Cumulative Distribution Function of Node Degrees



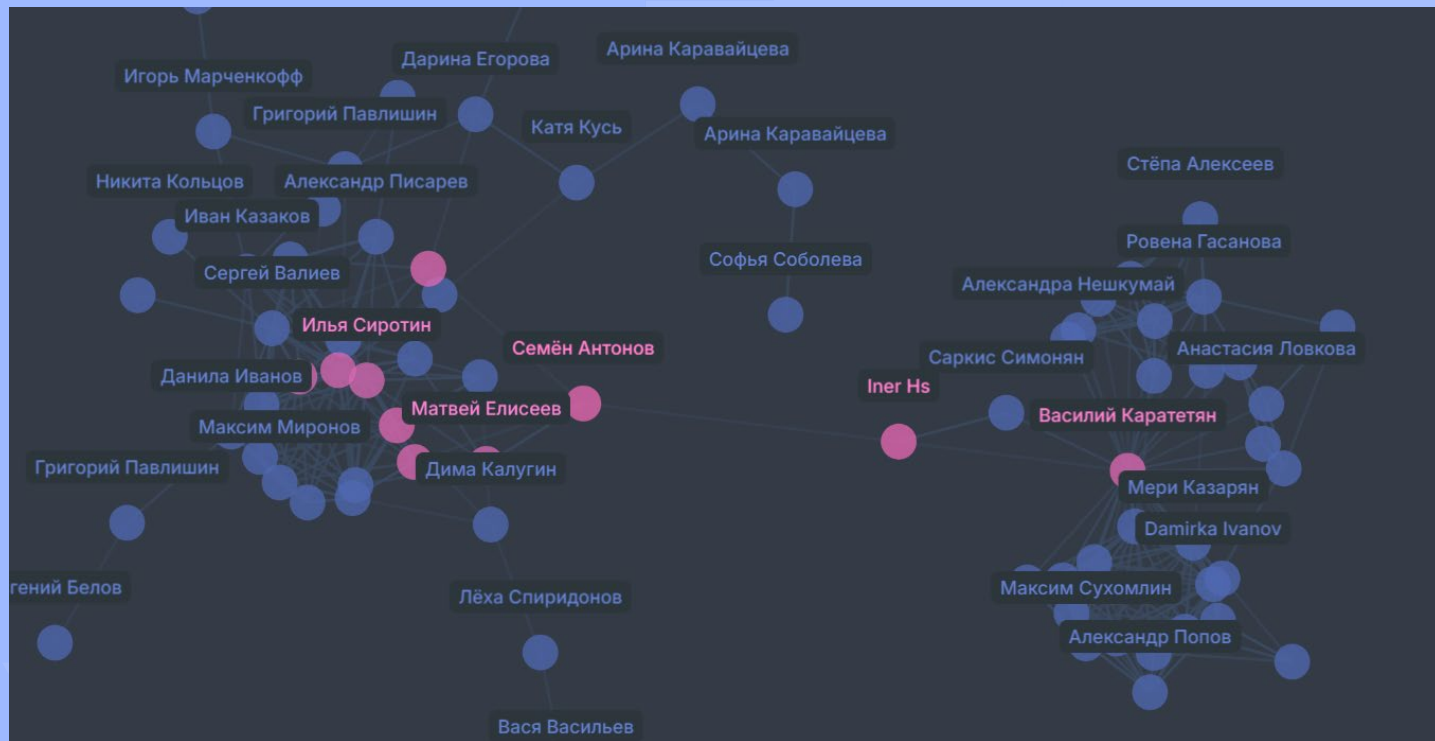


Top 10 nodes are highlighted

# Closeness centrality

Top 10 nodes are highlighted

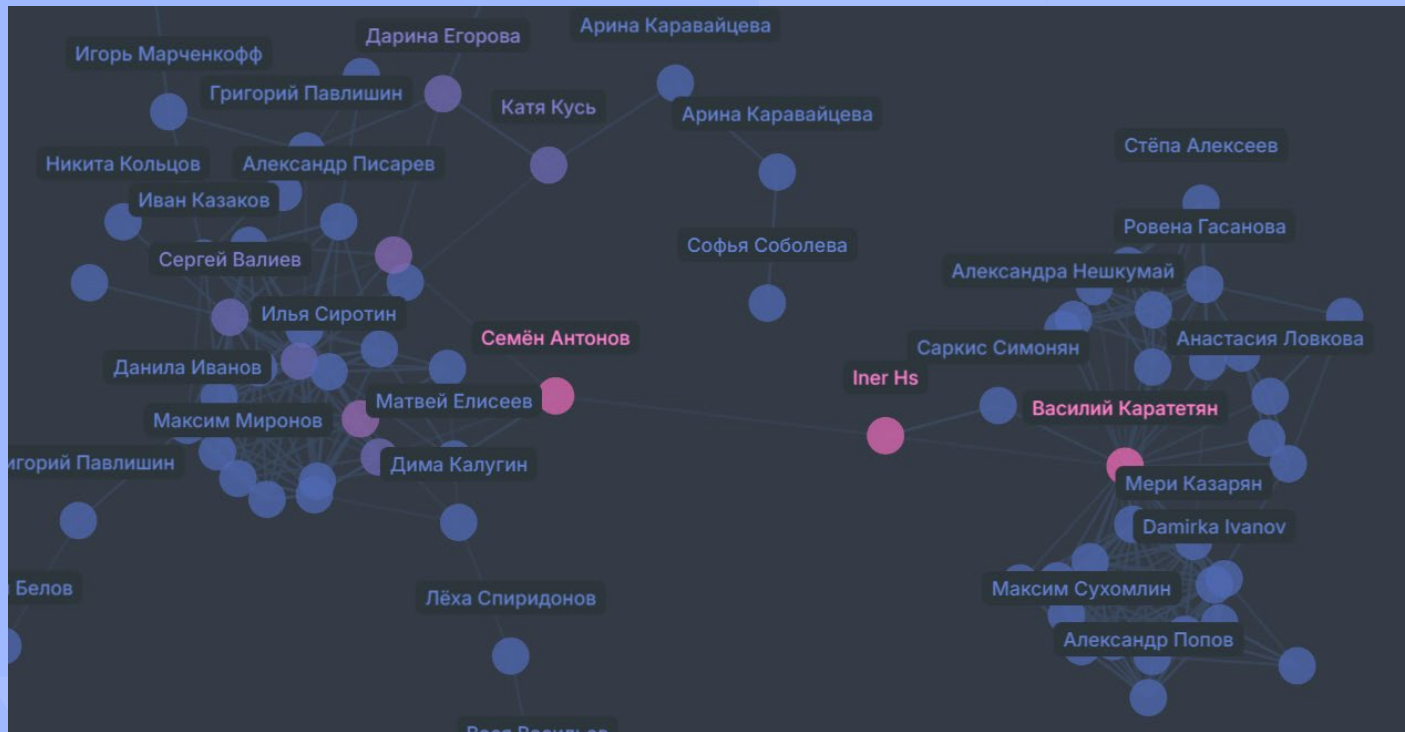
Name	Closeness centrality value
Семён Антонов	0.353
Iner Hs	0.338
Эльвин Гасанов	0.327
Горислав Егоров	0.323
Denis Antonov	0.320
Василий Каратетян	0.320



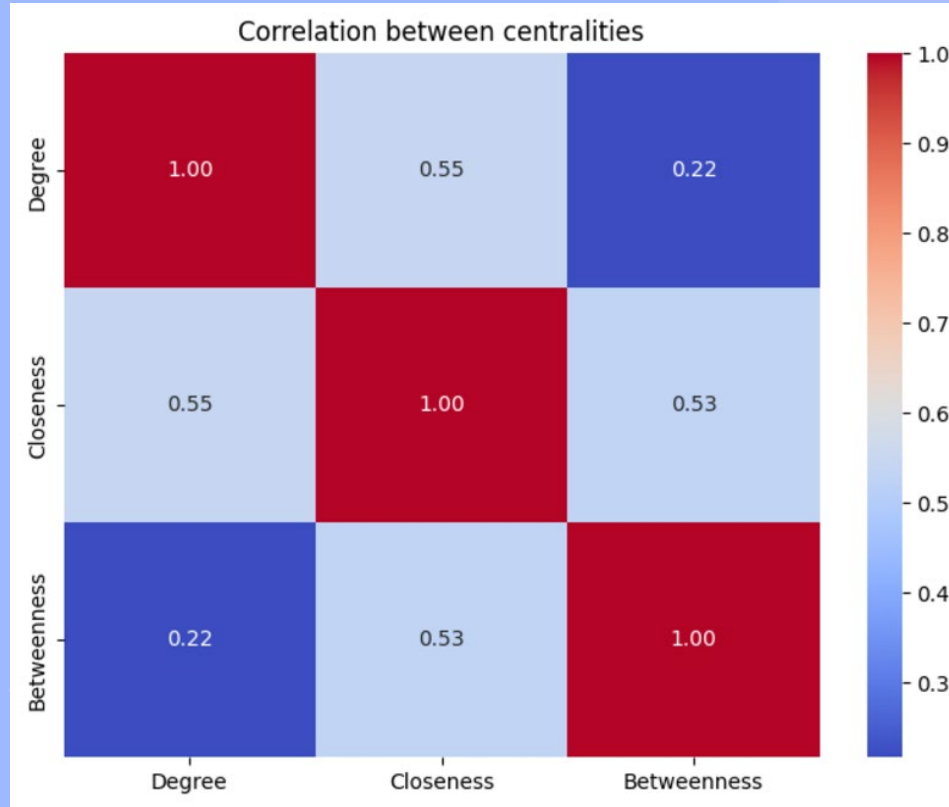
# Betweenness centrality

Top 10 nodes are highlighted

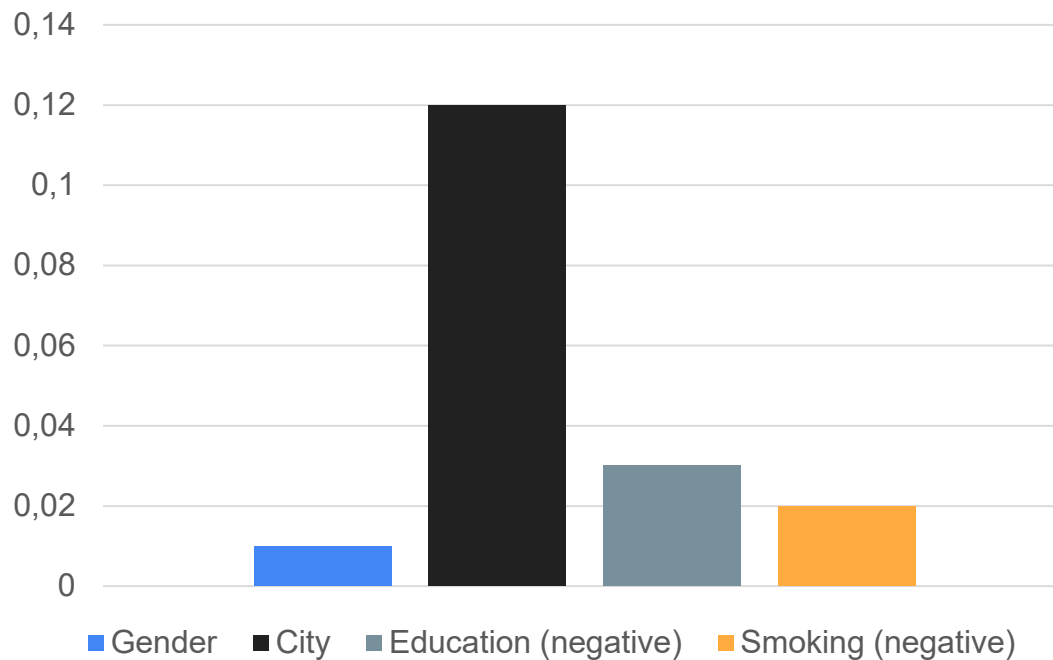
Name	Betweenness centrality value
Василий Каратетян	0.566
Семён Антонов	0.504
Iner Hs	0.497
Эльвин Гасанов	0.221
Горислав Егоров	0.162
Дарина Егорова	0.099



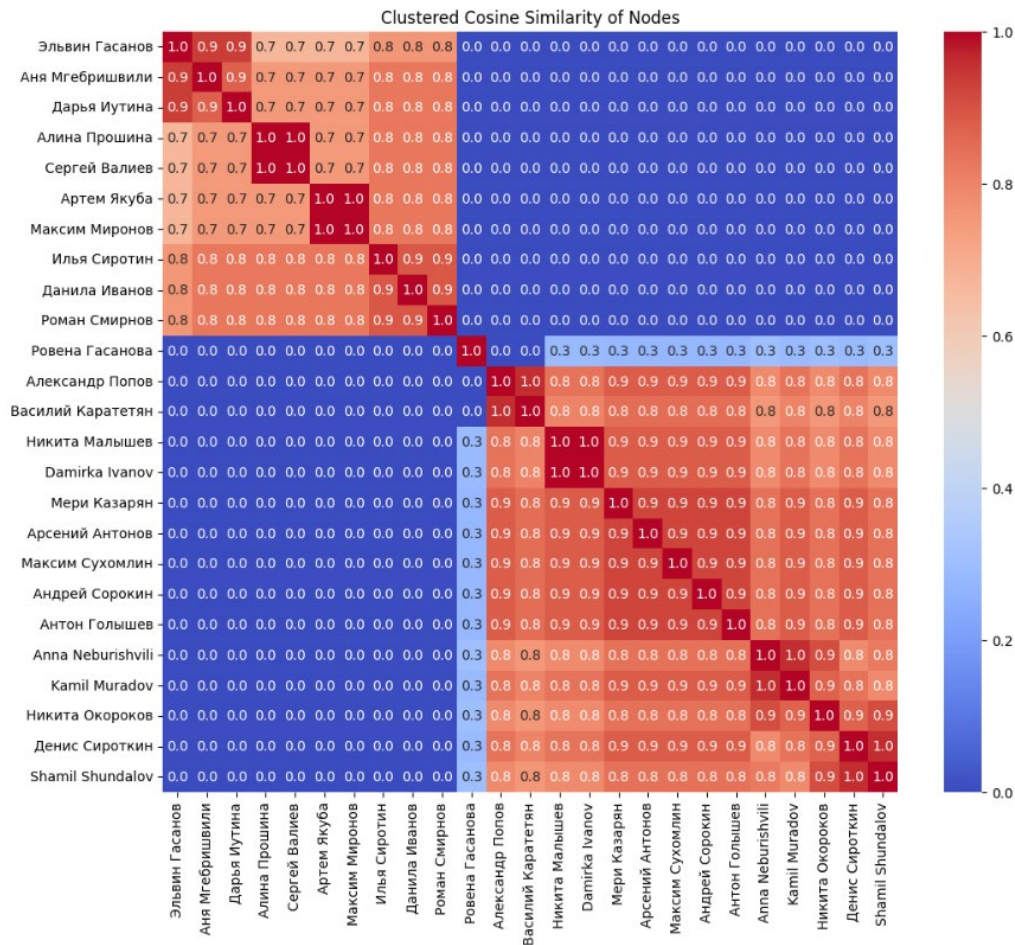
# Correlation between centralities



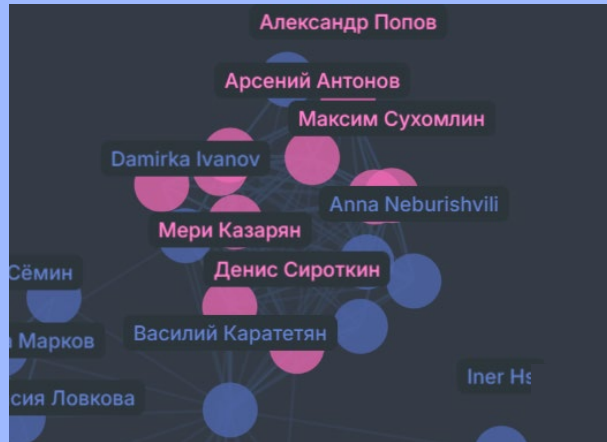
# Assortativity



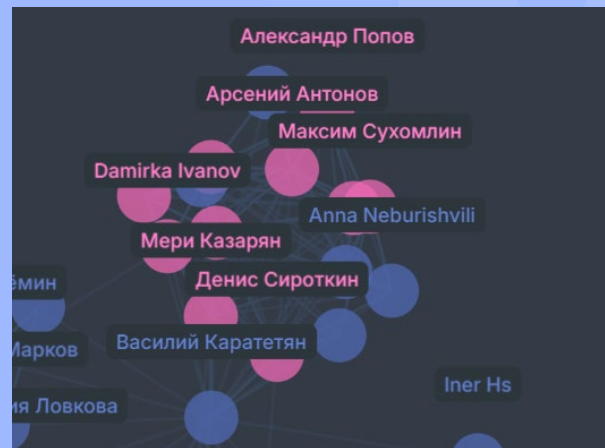
- Cosine similarity on adjacency matrix
- Hierarchical clustering



# Largest cliques



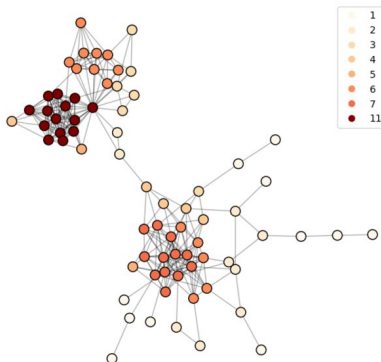
4 similar  
largest  
cliques with  
size = 10



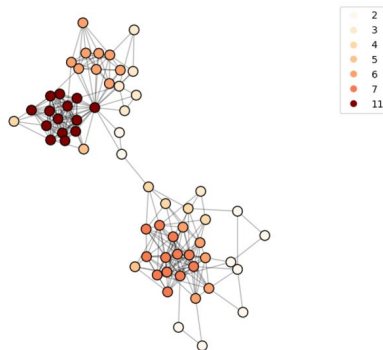


# K-cores decomposition

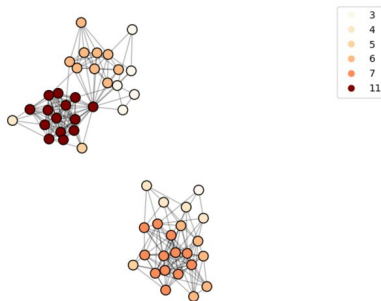
k-shells on 1-core



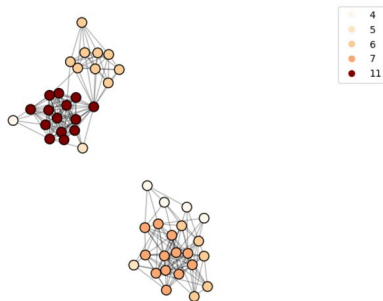
k-shells on 2-core



k-shells on 3-core



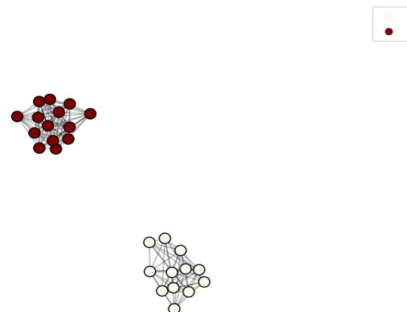
k-shells on 4-core



k-shells on 5-core



k-shells on 7-core



k-shells on 6-core

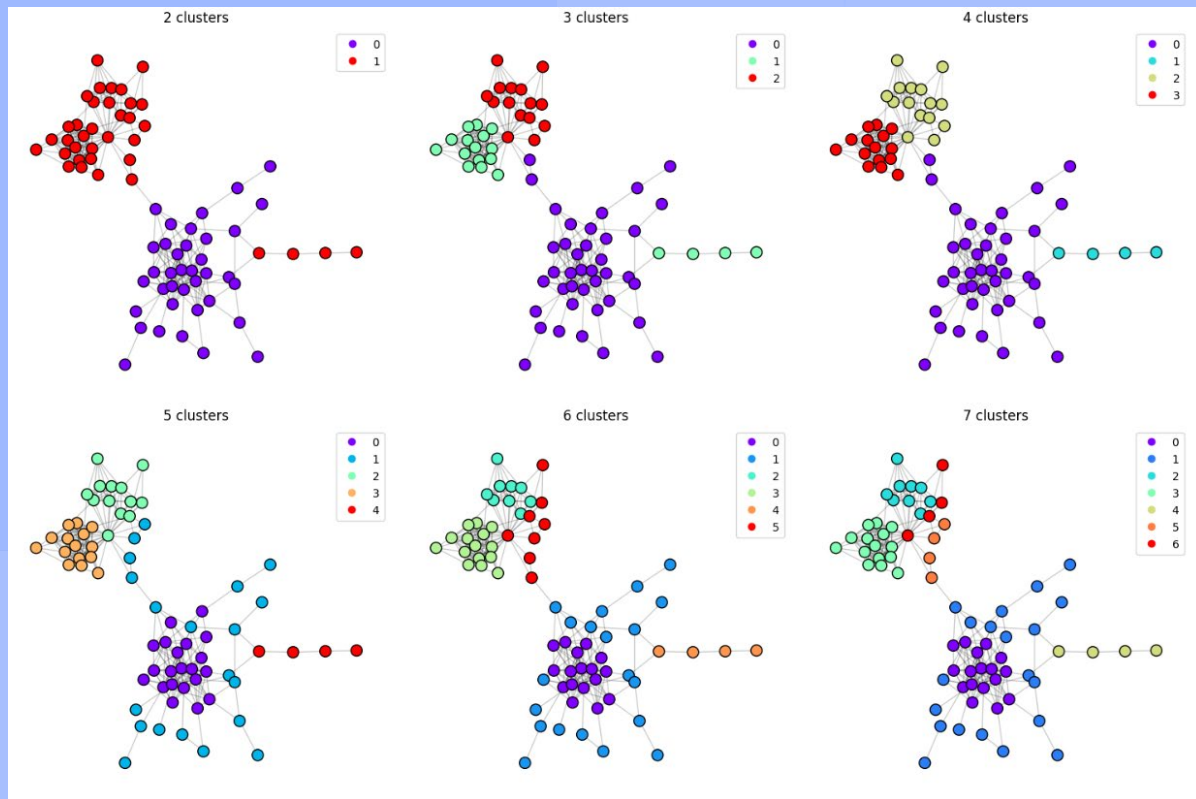
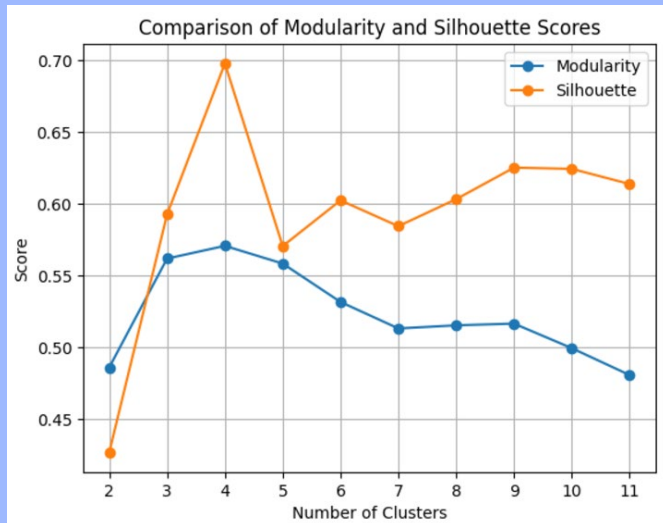


k-shells on 8-core



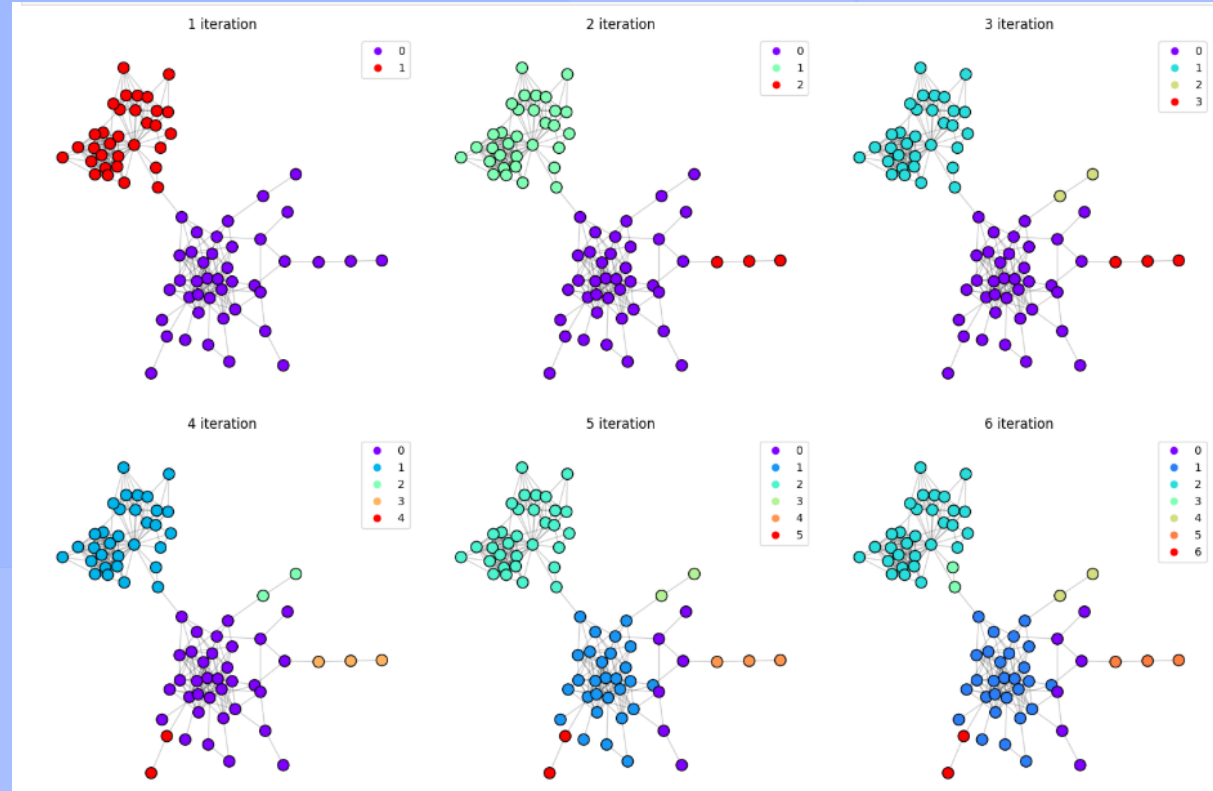
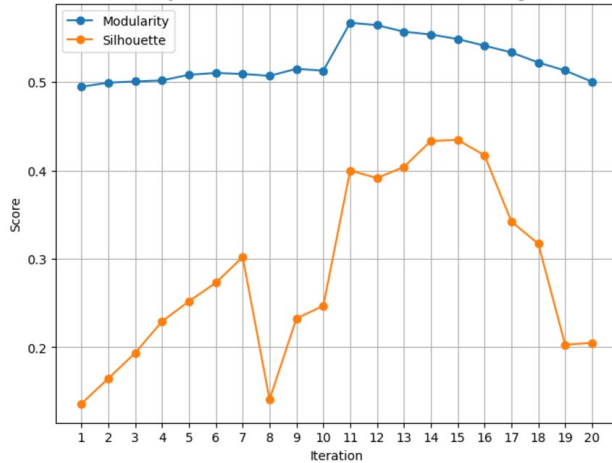


# Laplacian eigenmaps + K-means

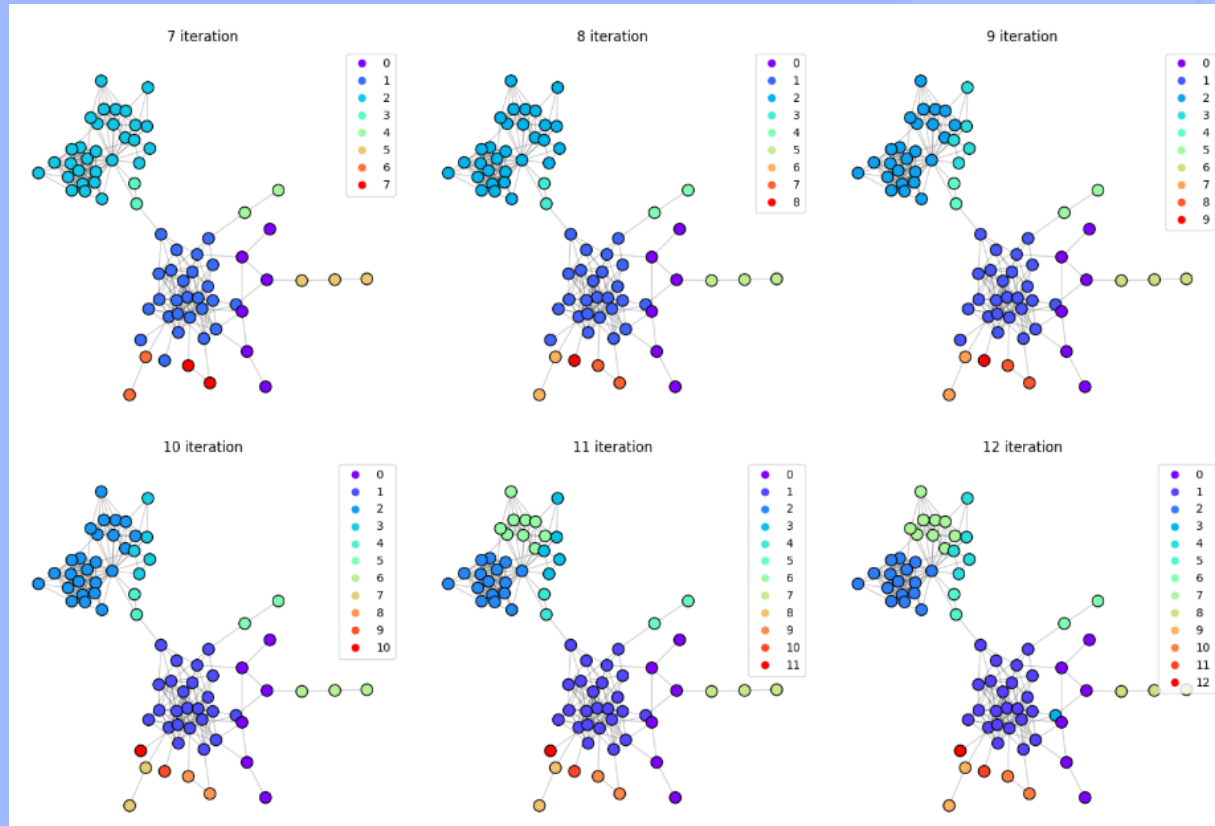


# Girvan - Newman

Modularity and Silhouette Scores for Girvan-Newman Algorithm



# Girvan - Newman



The End. Thank you!