

Tracking Refugee Movements



Chloe McGlynn, Melody Lamphear, Eric Ha, Sammy Cayo

Fundamentals of Data Engineering

Spring 2024 - Crook - Section 4

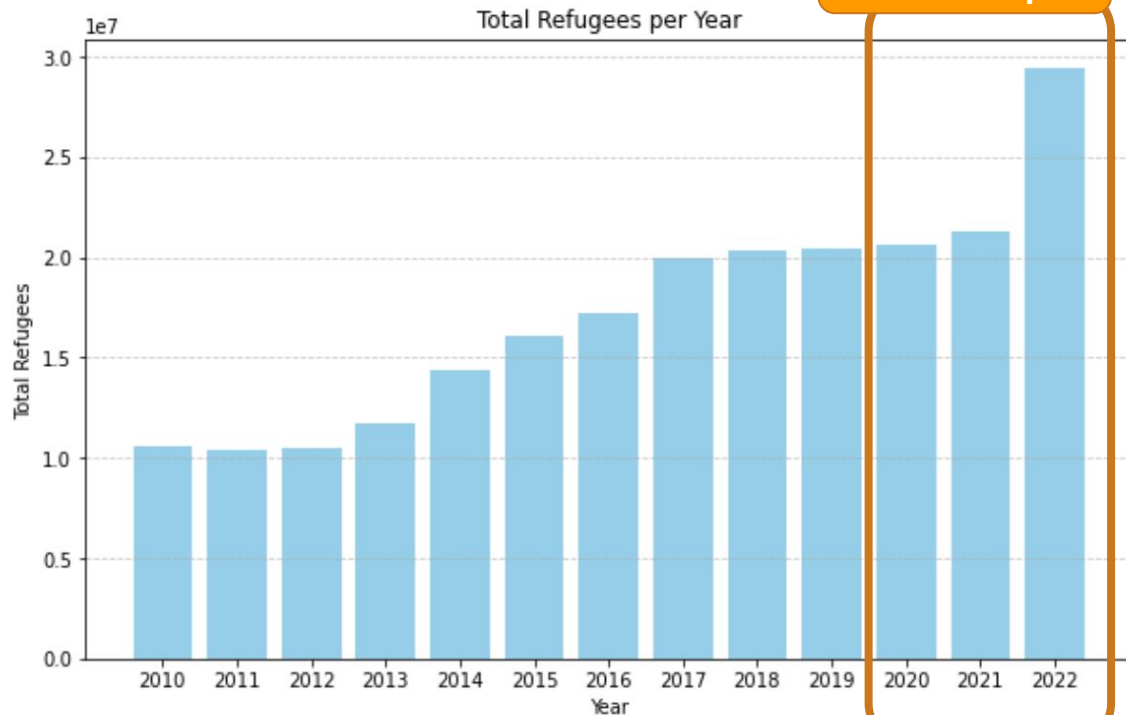
UC Berkeley MIDS Program

Background

- **Combined data from three major sources:**
 - United Nations High Commissioner for Refugees' (UNHCR)
 - United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA)
 - Internal Displacement Monitoring Centre (IDMC)
- **R package {refugees} includes 8 datasets total**
 - population.csv and solutions.csv
 - Country of Origin (coo), Country of Asylum (coa), Refugees, and Returned Refugees



Data Exploration



Origin (COO)



13,396,683



8,374,544



5,707,442



27.5M | 38%

Asylum (COA)



10,980,438



5,023,459



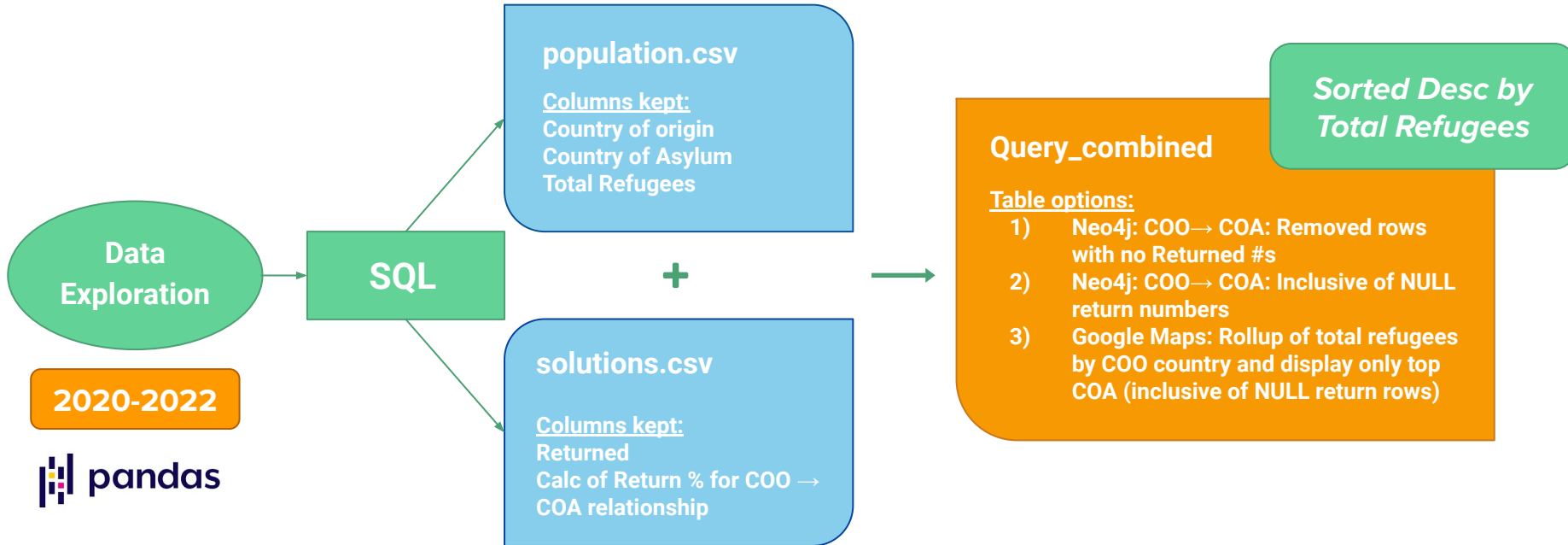
4,673,795



20.7M | 29%

Data Exploration

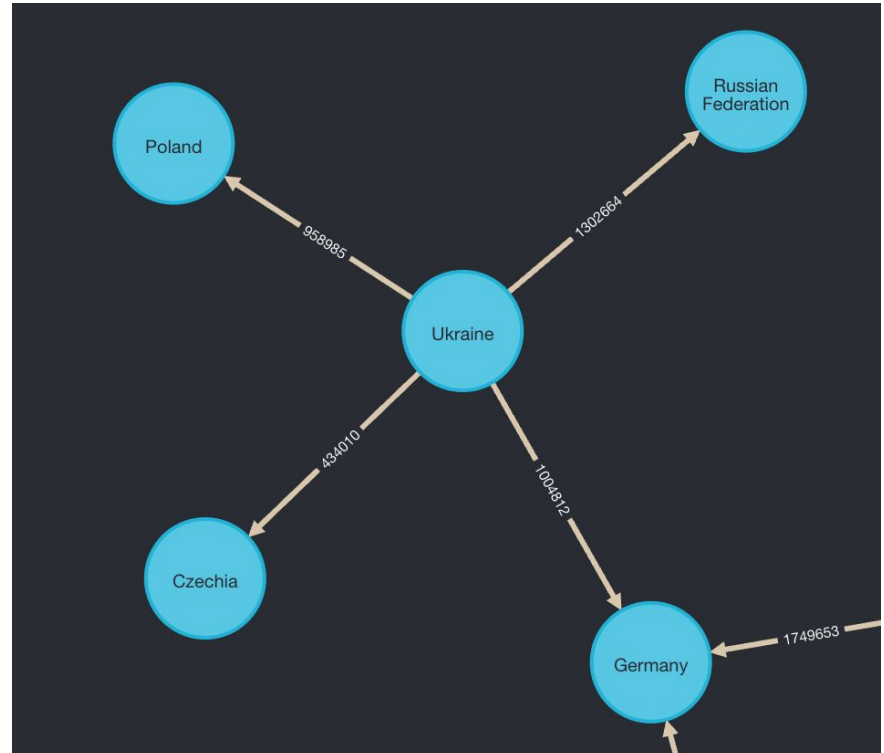
Research Question: Where are most refugees coming from/moving to, and how many were turned away from country of asylum during recent years (2020-2022)?



[illegible]

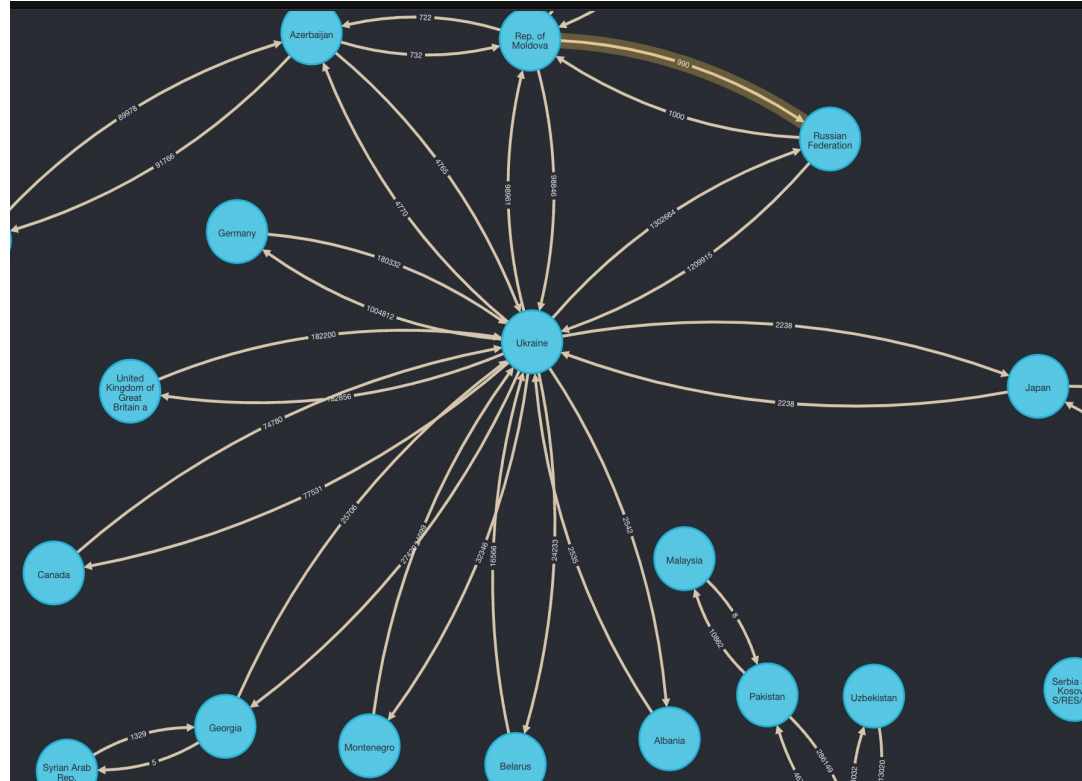
Neo4J Graph Example 1

- **Top 25 rows with NAs under “returners”**
 - Graph prioritizes COAs that received the highest number of refugees from respective COO
- **Graph Approach**
 - Each Country (COO or COA) is a node
 - One way relationship function for refugees leaving COO → COA
- **Example: Ukraine (2020-2022)**
 - No returners in data subset
 - Relational arrows represent total number of refugees going to COAs
 - Top COAs for Ukraine by # of refugees leaving were:
 - Russian Federation
 - Germany
 - Poland
 - Czechia



Neo4J Graph Example 2

- **Top 25 rows without NAs under “returners”**
 - Graph prioritizes COAs that tracked higher number of refugees in **and** at least some number of refugees returned to their COO
- **Graph Approach**
 - Each Country (COO or COA) is a node
 - One way relationship function for refugees leaving COO → COA
 - One way relationship function for refugees returning from COA → COO
- **Example: Ukraine (2020-2022)**
 - Returners present in data subset
 - Relational arrows represent total number of refugees leaving and returning
 - More countries included



Sample of Algorithm Results

- Influential node ranking

- PageRank (PR)
- Personalized PageRank (PPR)

- Community Detection

- Louvain Modularity

Top 50 rows without NAs

PR

name	page_rank
Rep. of Moldova	1.602180
Ukraine	1.334964
China, Hong Kong SAR	1.087721
Afghanistan	1.075422
Myanmar	1.035138

PPR: Moldova

name	page_rank
Rep. of Moldova	0.358620
Ukraine	0.093080
Azerbaijan	0.031350
Georgia	0.029174
Syrian Arab Rep.	0.028299

name	community	intermediate_community
Armenia	10	[11, 10]
Azerbaijan	10	[11, 10]
Bulgaria	10	[10, 10]
China	10	[10, 10]
Israel	10	[10, 10]

name	community	intermediate_community
Afghanistan	35	[1, 35]
Bangladesh	35	[35, 35]
China, Hong Kong SAR	35	[31, 35]
Iran (Islamic Rep. of)	35	[1, 35]
Malaysia	35	[35, 35]
Pakistan	35	[35, 35]

Other Database Use Cases



- Store precomputed paths for refugee movement based on historical data
- Store heuristics identified from migration patterns
- Use as web server for Refugee Data Visualization Dashboard



- Front-End Cache for Refugee Data Visualization
- Managing User Preferences and Customizations

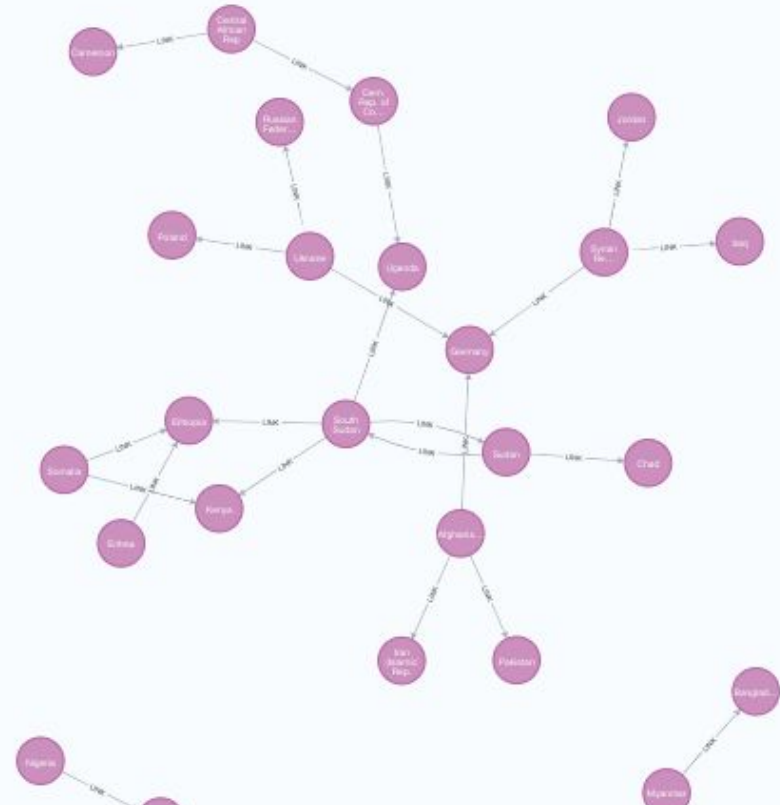
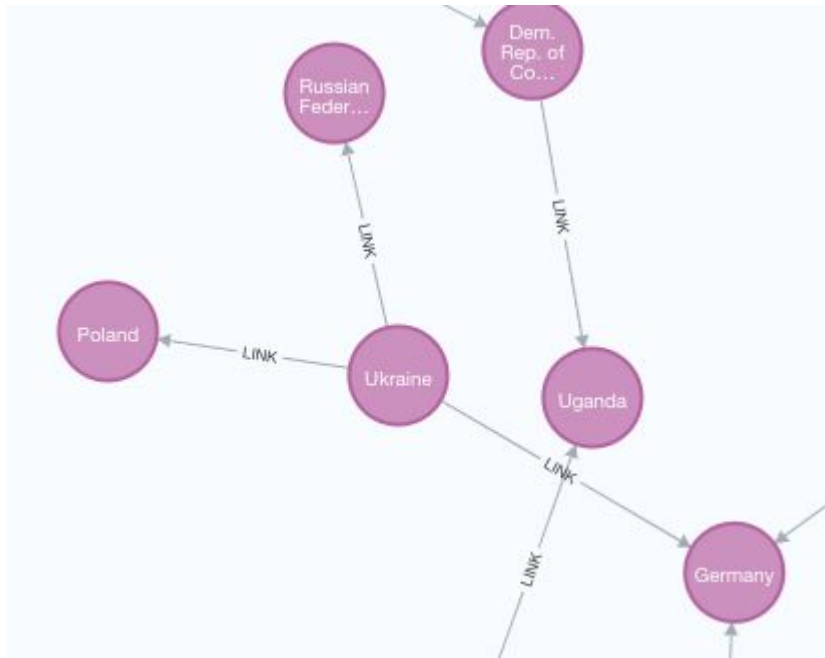
Conclusions

- **Dramatic increase in refugees due to war over 2020-2022**
 - Syria, Afghanistan and Ukraine creating the most refugees
- **Refugee count saw 10-year high in 2022**
 - ~30M refugees tracked in 2022
 - Prior decade average was only ~15M
- **Regional patterns of migration**
 - African refugees mostly settle in Africa
 - Everywhere else goes to Europe
 - EXCEPT for China, whose primary settlement country is the US

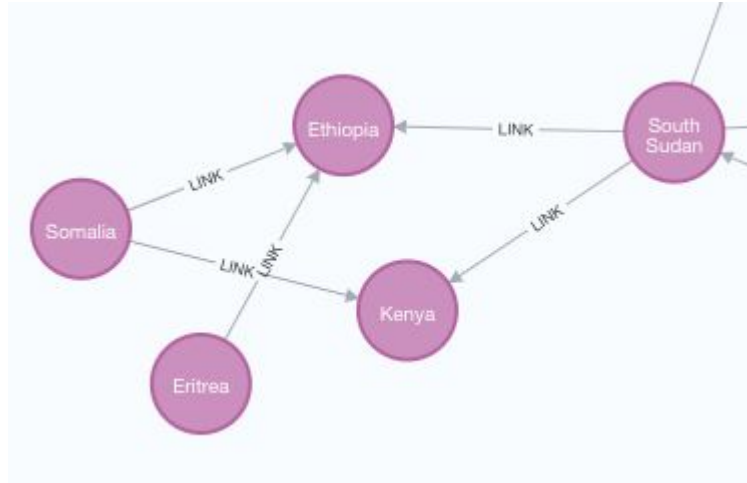
Questions?

Slides below are for team reference only

Top 25 with NAs (for drafting. Not presentation slide)

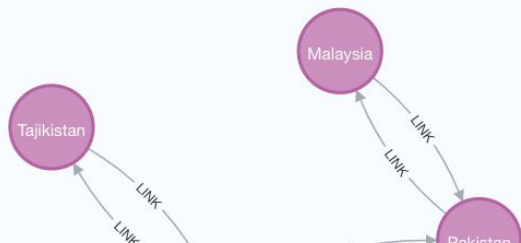


PageRank Top 25 with NAs (for drafting. Not presentation slide)

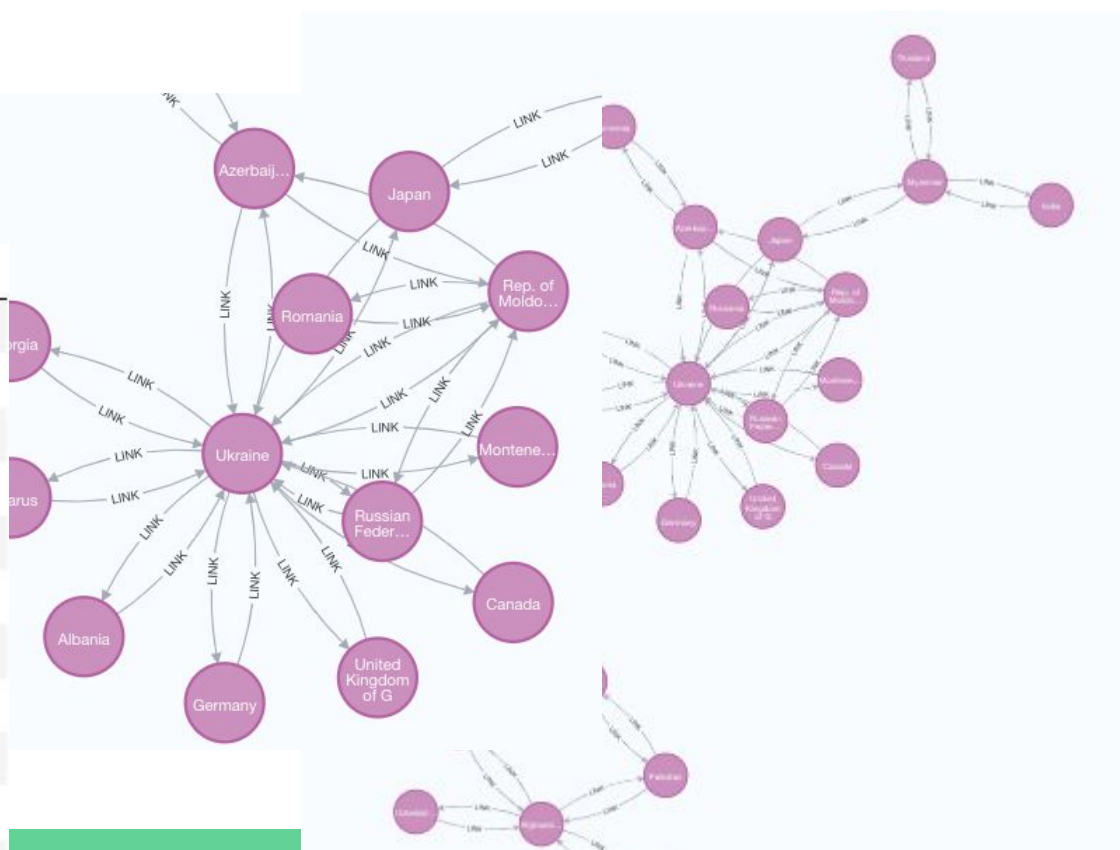


	name	page_rank
0	Ethiopia	1.033426
1	Uganda	1.010863
2	Bangladesh	0.997500
3	Niger	0.997500
4	Germany	0.989583
5	Kenya	0.985926
6	Chad	0.974054
7	South Sudan	0.974054
8	Cameroon	0.973750
9	Dem. Rep. of the Congo	0.973750
10	Iran (Islamic Rep. of)	0.965833

Top 25 without NAs (for drafting. Not presentation slide)



	name	page_rank
0	Ukraine	1.340553
1	Afghanistan	1.119919
2	Myanmar	1.071134
3	Rep. of Moldova	1.045434
4	Azerbaijan	1.017509
5	Pakistan	1.012765
6	Georgia	1.004849
7	North Macedonia	1.000000
8	Serbia and Kosovo: S/RES/1244 (1999)	1.000000
9	Malaysia	0.975319

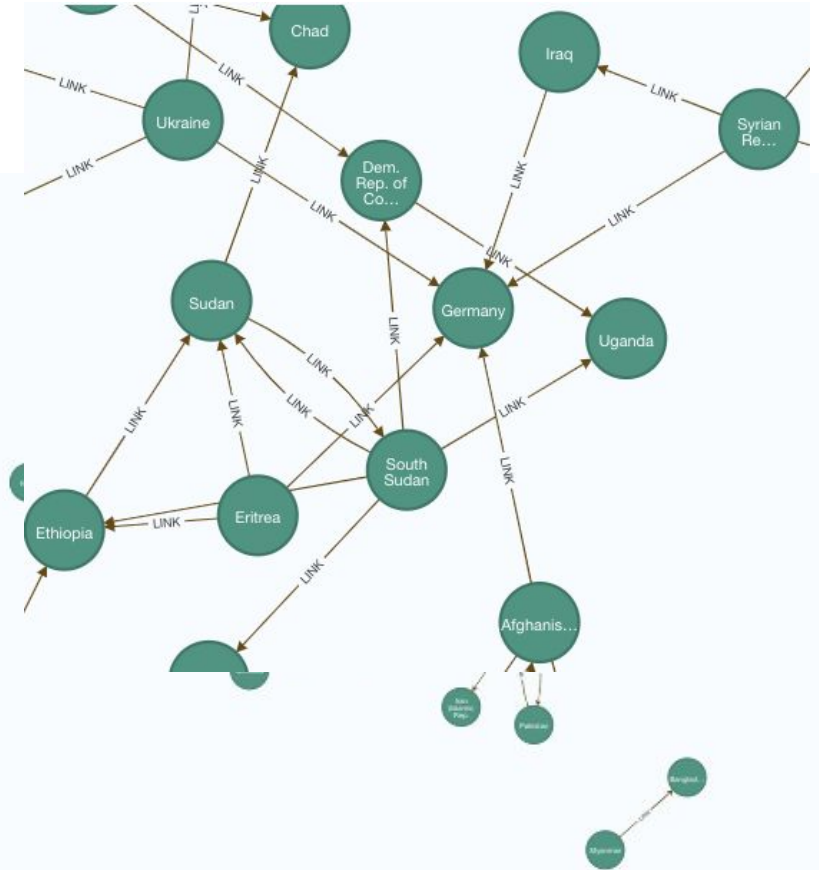


Personalized PageRank of Ukraine (for drafting)

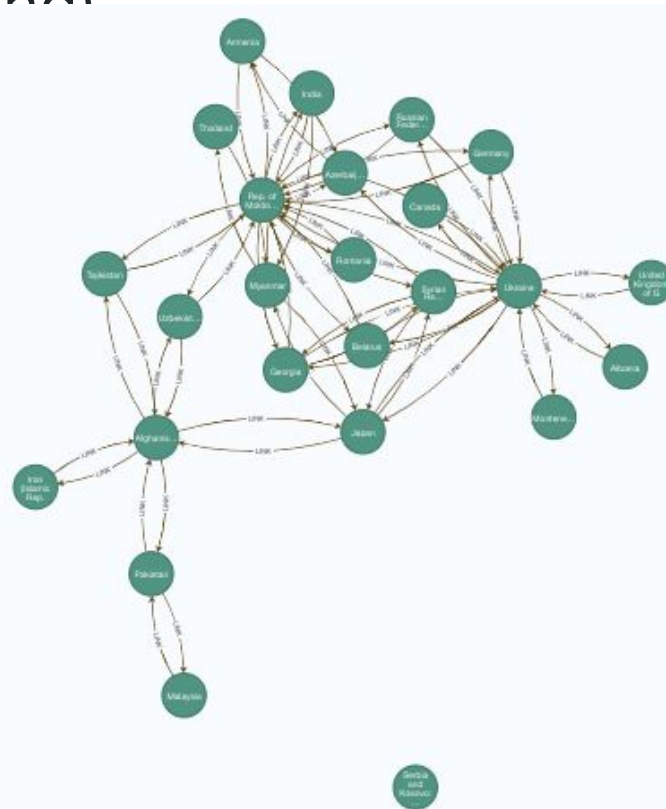
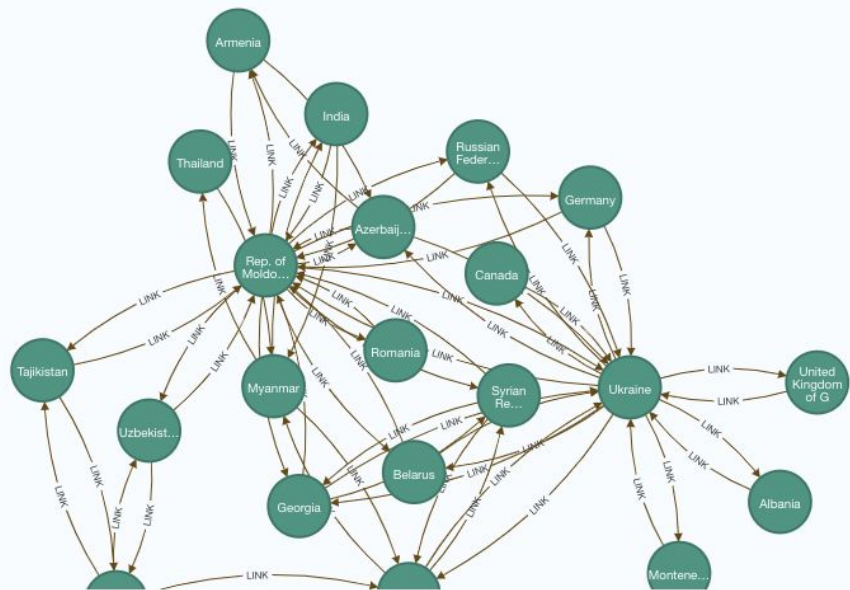
	name	page_rank
0	Ukraine	0.395704
1	Rep. of Moldova	0.082945
2	Azerbaijan	0.062765
3	Russian Federation	0.047934
4	Georgia	0.047254
5	Japan	0.039073
6	Myanmar	0.030622
7	Albania	0.030462
8	Belarus	0.030462
9	Canada	0.030462
10	Germany	0.030462

Top 50 With NAs (for drafting)

	name	page_rank
0	Germany	1.047785
1	Sudan	1.025161
2	Dem. Rep. of the Congo	1.023729
3	Nigeria	0.999541
4	Niger	0.998739
5	Afghanistan	0.998332
6	United Rep. of Tanzania	0.998140
7	Ecuador	0.997500
8	Spain	0.997500
9	United States of America	0.997500
10	Chad	0.991462



Top 50 Without NAs (for drafting)



0

1

2

3

4

5

6

7

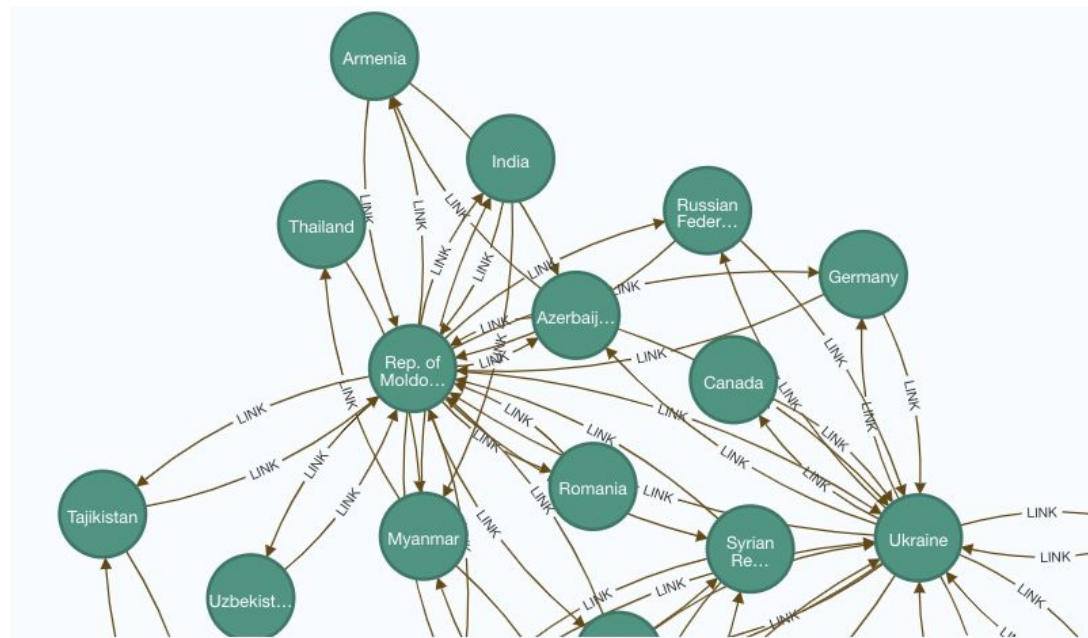
8

9

10

Personalized page rank of Rep of Moldova (for drafting)

	name	page_rank
0	Rep. of Moldova	0.358620
1	Ukraine	0.093080
2	Azerbaijan	0.031350
3	Georgia	0.029174
4	Syrian Arab Rep.	0.028299
5	Afghanistan	0.027159
6	Armenia	0.024000
7	Japan	0.023465
8	Belarus	0.021215
9	Germany	0.021215
10	Russian Federation	0.021215



Louvain Modularity - Top 50 w/o NAs (for drafting)

name	community	intermediate_community
Armenia	10	[11, 10]
Azerbaijan	10	[11, 10]
Bulgaria	10	[10, 10]
China	10	[10, 10]
Israel	10	[10, 10]

node	degree	closeness	betweenness	community
Afghanistan	5	0.4452	446	35
Albania	1	0.3768	0	25
Armenia	2	0.4184	0	10
Azerbaijan	3	0.4645	0	10
Bangladesh	2	0.2529	40	35
Belarus	2	0.4513	0	25

name	community	intermediate_community
Afghanistan	35	[1, 35]
Bangladesh	35	[35, 35]
China, Hong Kong SAR	35	[31, 35]
Iran (Islamic Rep. of)	35	[1, 35]
Malaysia	35	[35, 35]
Pakistan	35	[35, 35]
Rwanda	35	[31, 35]
Tajikistan	35	[1, 35]
Uzbekistan	35	[1, 35]
Yemen	35	[31, 35]