Migration patterns within the United States

Final Report CNT 5805

Team: 5

By
Dharitrikumari Rathod
Umapathy Beema
December 3, 2020

This year we have struggled a lot with the emergence of the new virus, Covid-19. People have suffered a lot of pain due to the loss of a family member, loss of a job, closed businesses, lay-offs, stay home orders, working from home, virtual classes, etc. During this time, our first thought was doing a project on migration, due to Covid-19, but we didn't have reliable data to study this problem so we decided to go with domestic migration patterns in the United States. Every year nearly 32 million people migrate from one place to another in the US for different reasons such as family, employment, housing, climate change, education attainment, etc.

1. **Purpose, Summary, and Research Questions:**
   - **1.1. Purpose:** It was very interesting to study the internal migration in the US and find its patterns. Also, to understand its causes and effects to educate ourselves and inform others regarding the current trend in migration patterns, which might give a good insight into the job market, real estate, etc.
   - **1.2. Summary**: To study the internal migration within the US, we have collected data from the census.gov[a][b][c] on the mobility and migration tab. For this project, we started with regional migration in the US then studied state to state migration (2018) followed by metro to metro migration (2014-2018). The Regional migration data suggested that there was a pattern in the regional migration. People in the northeast region are moving out of the region and southern regions are constantly gaining new populations. Then, our next question was to find the states in these regions with more inflows and the states with more outflows. Furthermore, we wanted to find the hubs in those states for inflow as well as outflows.

     After careful study of the regional migration, we found homophily in the regional migration. People in the west tend to live in the west and people on the east coast tend to live on the east coast. Also for this migration, the selection of a state is random, but selecting a metropolitan area within the selected state follows the power-law distribution.

     Once we found the pattern, we wanted to rank the states with the highest net in-migration and the state with the highest net out-migration. To do this, first, we ranked all the states with the highest in-migration and found the most contributors to those states. Then we ranked all the states with the highest out-migration, but we found many common states between in-migration and out-migration. Hence, we decided to go with net migration (net migration= in-migration - out-migration) in the state. Moreover, we found top recipients of the states with the highest out-migration.

     Once we have listed the top 10 states with net migration, our next task was to find the metropolitan areas in those states which attract more people. Therefore, we found and ranked the top 5 metro areas in each of the top 5 states with the highest net in-migration and found top hubs in those states. Similarly, we listed top hubs for the states with the highest net out-migration. Furthermore, we found the most popular destination in the US.

     Lastly, we tried to understand the reasons to move from one state to another state based on the data from census.gov. Data for reasons to move was regional and overall US migration reasons. This data shows that every year 3.5 million people move due to

employment and 9 million people move due to housing-related reasons. Also, every year approximately 400,000 people move due to climate change and global warming. Research shows that global warming will affect the migration pattern in the US and worldwide significantly in this century [3]. For the effects, an exodus to any state or region is bad for their economy. On the other hand, states or metropolitan areas with high inflows will have a lot of opportunities for real estate, labor market, infrastructure, etc. Due to the migration towards western and southern states, especially swing states will become more important to decide white house administration as we have seen in the year 2020.

1.3. Research question: In this study, we tried to answer the following questions:

a) What is the domestic migration pattern in the United States?
b) Where people are moving?
c) What are the causes and effects of domestic migration in the US?

2. Network Overview Statistics and Data laboratory:

2.1. Network overview: The network for migration will have states/metro as nodes and if there is a migration from one state to another then there is an edge between those two states/metros. We have studied two networks for this analysis: one for the state to state migration and Metro to metro migration. In the state to state migration network, there are 52 states including DC and Puerto Rico, and 2409 edges. In the metro to metro migration network, we have 390 nodes and 70340 edges. As the data file was too big for Gephi, we have made individual networks for the top 5 states with in-migration and out-migration using Excel and Gephi. All the networks are directed and weighted networks.

2.2. State to State Data Laboratory: State to state Migration -2018

a) Node table:

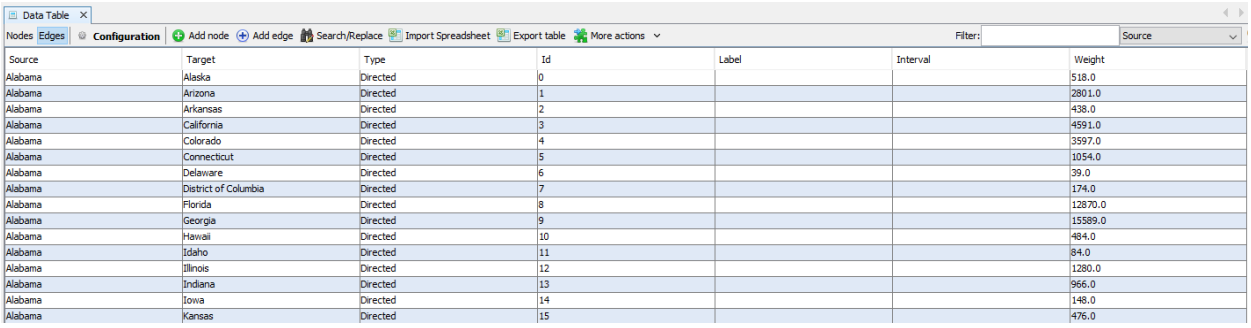
Data Table

Nodes Edges | Configuration | Add node | Add edge | Search/Replace | Import Spreadsheet | Export table | More actions | Filter: | Id

Id	Label	In-Deg...	Out-D...	Degree	Weighted I...	Weighted O...	Weighted...	Eccent...	Closeness ...	Harmonic Closene...	Betweenness...	Authority	Hub	Modulari...	Clustering C...	Strongly-Con...	Clustering C...	Eigenvector ...	net weighte...
Alabama	Alabama	43	49	92	110258.0	98194.0	208452.0	2.0	0.962264	0.980392	3.571517	0.129622	0.146003	1	0.964898	0	0.910612	0.859438	12064
Alaska	Alaska	47	43	90	32443.0	39246.0	71689.0	2.0	0.864407	0.921569	2.608181	0.140561	0.130803	0	0.97449	0	0.92432	0.932472	-6803
Arizona	Arizona	51	51	102	275003.0	179781.0	454784.0	1.0	1.0	1.0	7.553601	0.150042	0.149863	0	0.964706	0	0.904706	1.0	95222
Arkansas	Arkansas	46	42	88	74872.0	69724.0	144596.0	2.0	0.85	0.911765	2.411335	0.137996	0.127845	3	0.977837	0	0.929521	0.913698	5148
California	California	51	51	102	502273.0	691665.0	1193938.0	1.0	1.0	1.0	7.553601	0.150042	0.149863	0	0.964706	0	0.904706	1.0	-189392
Colorado	Colorado	51	50	101	240257.0	204182.0	444439.0	2.0	0.980769	0.990196	7.08498	0.150091	0.147576	3	0.964706	0	0.905098	1.0	36075
Connect...	Connect...	44	46	90	89203.0	115243.0	204446.0	2.0	0.910714	0.95098	4.451033	0.131373	0.136477	2	0.972518	0	0.921986	0.871806	-26040
Delaware	Delaware	36	34	70	34636.0	33395.0	68031.0	2.0	0.75	0.833333	1.618118	0.107095	0.104024	1	0.971207	0	0.931894	0.708476	1241
District ...	District ...	41	43	84	47738.0	60363.0	108101.0	2.0	0.864407	0.921569	3.054475	0.122501	0.129364	1	0.971323	0	0.922757	0.813048	-12625
Florida	Florida	51	51	102	631971.0	476893.0	1108864.0	1.0	1.0	1.0	7.553601	0.150042	0.149863	1	0.964706	0	0.904706	1.0	155078
Georgia	Georgia	51	51	102	275655.0	231510.0	507165.0	1.0	1.0	1.0	7.553601	0.150042	0.149863	1	0.964706	0	0.904706	1.0	44145
Hawaii	Hawaii	45	47	92	54142.0	67505.0	121647.0	2.0	0.927273	0.960784	3.649532	0.135749	0.140454	0	0.971429	0	0.917143	0.896826	-13363
Idaho	Idaho	43	40	83	80468.0	55183.0	135651.0	2.0	0.822581	0.892157	1.98784	0.129802	0.121488	0	0.975177	0	0.930408	0.86222	25285
Illinois	Illinois	50	50	100	209297.0	305627.0	514924.0	2.0	0.980769	0.990196	7.035645	0.147405	0.147229	1	0.964706	0	0.90549	0.981572	-96330
Indiana	Indiana	48	48	96	147817.0	136563.0	284380.0	2.0	0.944444	0.970588	4.547457	0.142885	0.143259	1	0.971429	0	0.91551	0.951901	11254
Iowa	Iowa	45	47	92	74832.0	70659.0	145491.0	2.0	0.927273	0.960784	3.014041	0.135737	0.141012	3	0.97534	0	0.922619	0.897774	4173
Kansas	Kansas	45	47	92	95900.0	96850.0	192750.0	2.0	0.927273	0.960784	2.799518	0.135957	0.141126	3	0.97534	0	0.923895	0.899466	-950
Kentucky	Kentucky	48	49	97	111578.0	100514.0	212092.0	2.0	0.962264	0.980392	4.771279	0.142968	0.145786	1	0.964706	0	0.906667	0.948873	11064
Louisiana	Louisiana	47	47	94	81593.0	95326.0	176919.0	2.0	0.927273	0.960784	3.481876	0.140873	0.14099	3	0.97449	0	0.921769	0.932807	-13733
Maryland	Maryland	51	48	99	165178.0	182525.0	347703.0	2.0	0.944444	0.970588	7.071238	0.150228	0.141144	1	0.964706	0	0.905882	1.0	-17347
Massac...	Massac...	50	48	98	149719.0	180806.0	330525.0	2.0	0.944444	0.970588	6.239675	0.147704	0.142021	2	0.966531	0	0.910204	0.982336	-31087
Michigan	Michigan	51	50	101	139673.0	152605.0	292278.0	2.0	0.980769	0.990196	6.234176	0.150074	0.148357	1	0.964706	0	0.905098	1.0	-12932
Minnesota	Minnesota	49	48	97	100370.0	112228.0	212598.0	2.0	0.944444	0.970588	4.408305	0.145542	0.14391	3	0.969796	0	0.915102	0.968275	-11858
Mississippi	Mississippi	43	44	87	68303.0	76567.0	144870.0	2.0	0.87931	0.931373	2.201297	0.130339	0.132882	3	0.980574	0	0.933858	0.863885	-8264
Missouri	Missouri	50	49	99	153103.0	174949.0	328052.0	2.0	0.962264	0.980392	5.344625	0.147775	0.146119	3	0.964706	0	0.905882	0.982499	-21846
Montana	Montana	45	43	88	43506.0	35292.0	78798.0	2.0	0.864407	0.921569	2.664163	0.134979	0.130174	0	0.981383	0	0.933954	0.897332	8214
Nebraska	Nebraska	48	45	93	54619.0	54462.0	109081.0	2.0	0.894737	0.941176	3.638036	0.14353	0.134993	3	0.967347	0	0.913061	0.95032	157
Nevada	Nevada	48	50	98	127926.0	89876.0	217802.0	2.0	0.980769	0.990196	6.003727	0.142483	0.147738	0	0.964706	0	0.906274	0.947096	38050
New Ha...	New Ha...	42	37	79	49408.0	43793.0	93201.0	2.0	0.784615	0.862745	3.211219	0.126041	0.10993	2	0.972947	0	0.923671	0.833281	5615
New Jer...	New Jer...	45	50	95	165709.0	233079.0	398788.0	2.0	0.980769	0.990196	5.270996	0.134177	0.147802	2	0.968163	0	0.913469	0.890361	-67370
New Me...	New Me...	47	46	93	60793.0	74459.0	135252.0	2.0	0.910714	0.95098	4.122845	0.140234	0.1375	3	0.964706	0	0.908235	0.934687	-13666

Figure 2.0

b) Edges Table:



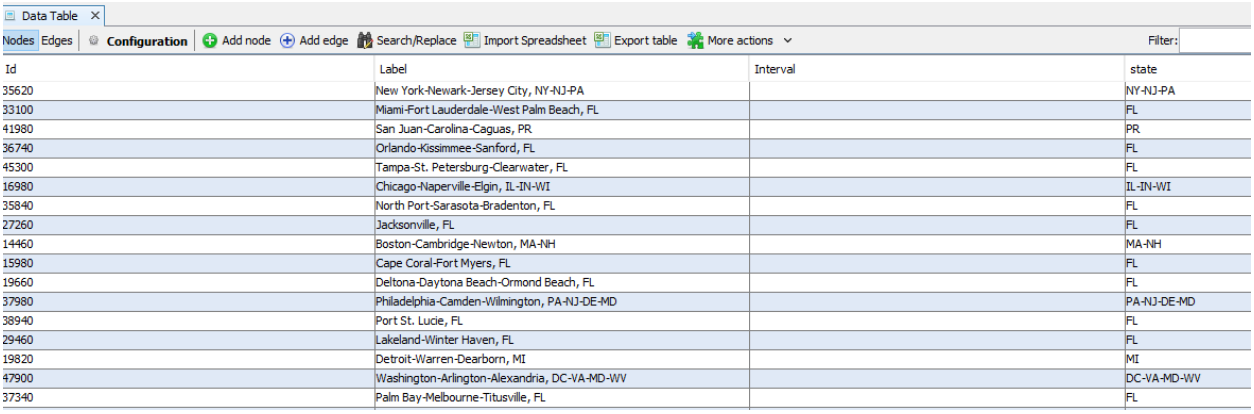
Data Table

Nodes Edges | Configuration | Add node | Add edge | Search/Replace | Import Spreadsheet | Export table | More actions | Filter: | Source

Source	Target	Type	Id	Label	Interval	Weight
Alabama	Alaska	Directed	0			518.0
Alabama	Arizona	Directed	1			2801.0
Alabama	Arkansas	Directed	2			438.0
Alabama	California	Directed	3			4591.0
Alabama	Colorado	Directed	4			3597.0
Alabama	Connecticut	Directed	5			1054.0
Alabama	Delaware	Directed	6			39.0
Alabama	District of Columbia	Directed	7			174.0
Alabama	Florida	Directed	8			12870.0
Alabama	Georgia	Directed	9			15589.0
Alabama	Hawaii	Directed	10			484.0
Alabama	Idaho	Directed	11			84.0
Alabama	Illinois	Directed	12			1280.0
Alabama	Indiana	Directed	13			966.0
Alabama	Iowa	Directed	14			148.0
Alabama	Kansas	Directed	15			476.0

2.3.Metro to Metro Data Laboratory: Metro – Metro migration 2014-2018

a) Node table:

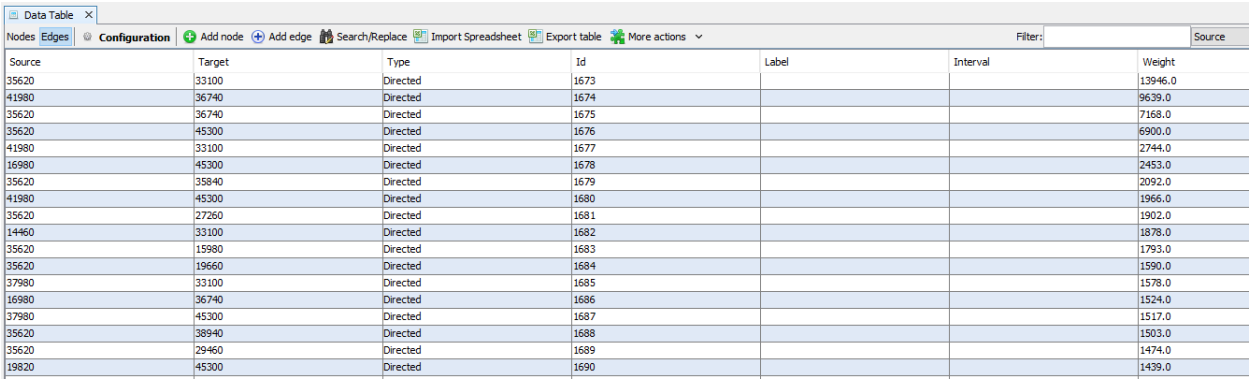


Data Table

Nodes Edges | Configuration | Add node | Add edge | Search/Replace | Import Spreadsheet | Export table | More actions | Filter:

Id	Label	Interval	state
35620	New York-Newark-Jersey City, NY-NJ-PA		NY-NJ-PA
33100	Miami-Fort Lauderdale-West Palm Beach, FL		FL
41980	San Juan-Carolina-Caguas, PR		PR
36740	Orlando-Kissimmee-Sanford, FL		FL
45300	Tampa-St. Petersburg-Clearwater, FL		FL
16980	Chicago-Naperville-Elgin, IL-IN-WI		IL-IN-WI
35840	North Port-Sarasota-Bradenton, FL		FL
27260	Jacksonville, FL		FL
14460	Boston-Cambridge-Newton, MA-NH		MA-NH
15980	Cape Coral-Fort Myers, FL		FL
19660	Deltona-Daytona Beach-Ormond Beach, FL		FL
37980	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD		PA-NJ-DE-MD
38940	Port St. Lucie, FL		FL
29460	Lakeland-Winter Haven, FL		FL
19820	Detroit-Warren-Dearborn, MI		MI
47900	Washington-Arlington-Alexandria, DC-VA-MD-WV		DC-VA-MD-WV
37340	Palm Bay-Melbourne-Titusville, FL		FL

b) Edge table:



Data Table

Nodes Edges | Configuration | Add node | Add edge | Search/Replace | Import Spreadsheet | Export table | More actions | Filter: | Source

Source	Target	Type	Id	Label	Interval	Weight
35620	33100	Directed	1673			13946.0
41980	36740	Directed	1674			9639.0
35620	36740	Directed	1675			7168.0
35620	45300	Directed	1676			6900.0
41980	33100	Directed	1677			2744.0
16980	45300	Directed	1678			2453.0
35620	35840	Directed	1679			2092.0
41980	45300	Directed	1680			1966.0
35620	27260	Directed	1681			1902.0
14460	33100	Directed	1682			1878.0
35620	15980	Directed	1683			1793.0
35620	19660	Directed	1684			1590.0
37980	33100	Directed	1685			1578.0
16980	36740	Directed	1686			1524.0
37980	45300	Directed	1687			1517.0
35620	38940	Directed	1688			1503.0
35620	29460	Directed	1689			1474.0
19820	45300	Directed	1690			1439.0

2.4.Statistics: we have used Gephi to analyze the graph. After loading the data in Gephi. First, we run the Gephi statistics for the state to state migration is as shown in Figure 2.1.

a) Average Degree: The average degree of the graph is 46.327. It is State to state migration; there are some states like Alaska, Hawaii, and Puerto Rico where some edges missing.

b) Avg. Weighted Degree: This statistics is very important for our analysis and we have heavily used these statistics on each network or graph of the state and metro network. Furthermore, Due to the directed graph, these statistics will generate two more

statistics for the network: weighted in-degree and weighted out-degree as shown in figure 2.0. Furthermore, using in and out weighted degrees we have calculated the net average degree for each node, where net weighted degree= weighted in-degree-weighted out-degree. For our analysis, the weighted net degree represents the net migration in the respective state.

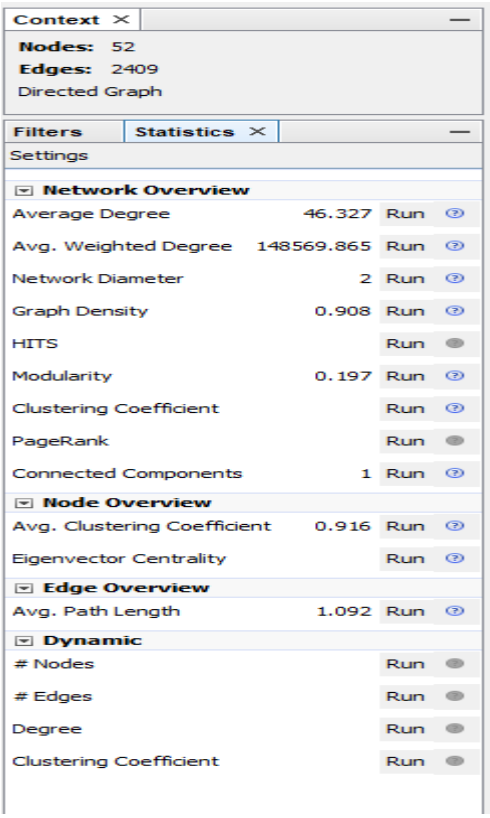


Figure 2.1

c) **Network diameter:** Network diameter is 2 for the state to state migration network as some of the states are not connected by an edge. Therefore, the shortest path between any two states is 2.

d) **Graph Density:** The graph density of the network is 0.908 which means a 90% dense graph as most of the states have direct edges to other states.

e) **Modularity:** This is the other most important statistics for this analysis as it helped to understand and identify the pattern in the state to state migration as well as metro to metro migration. How people move across the country such as regionwise, or neighboring states, one coast to the other, or stays on the same coast.

f) **Average clustering coefficient:** As it is an almost complete graph, neighbors of each node are also highly connected. Therefore, the average clustering coefficient is 0.916. In other words, approximately 92% of the neighbors of a particular node are connected via a direct edge.

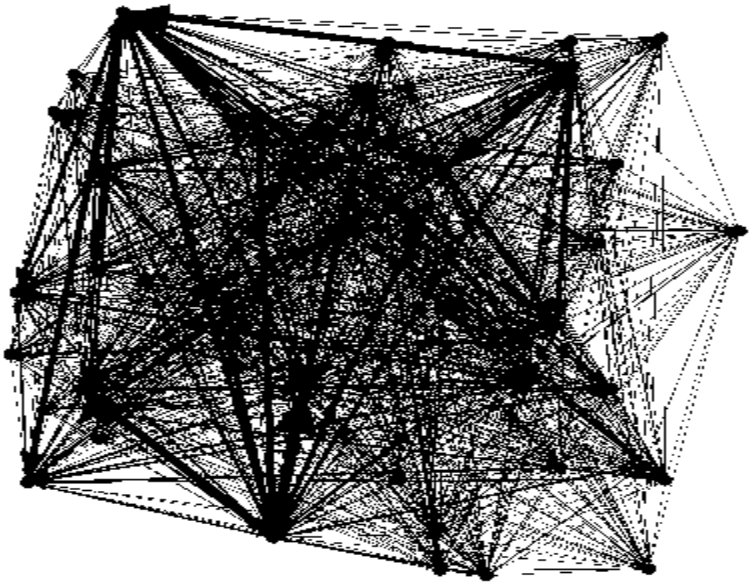
g) **Connected component:** This network has only 1 connected component as it is every node is connected to almost every other node. Hence, the average path length is 1.092.

h) **Smallworld and six degrees of distribution**: As the path length is 1.092 which is less than 6, the network represents the small world phenomenon.

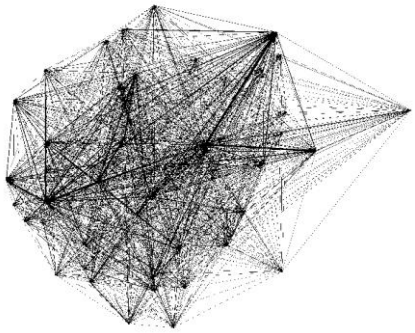
i) **Degree Distribution**: The Degree for the state to state migration is random, but metro to metro migration within the state follows the power-law distribution.

2.5.Initial Graph: From the following initial graph, we can ascertain that it is a connected, dense, directed, and weighted graph. Also, some of the nodes have high in-degrees and some of them have high out-degrees based on the size and direction of the arrows. Also, some nodes are with higher weighted edges than others based on the thickness of the edge. Besides, some of the nodes have don't seems to have very high in-degrees or out-

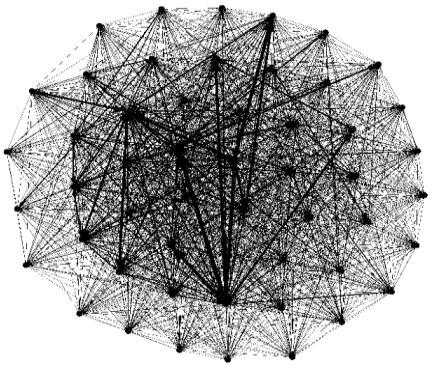
degrees. From the following graph, it will be interesting to study the nodes with high mobility.



3. **Layout Algorithms:** To study the network of the domestic migration in the United States, we have run 5 layouts as follows: Fruchterman Reingold, Yifan Hu, Force Altas, The Circular layout, and Radial axis layout.

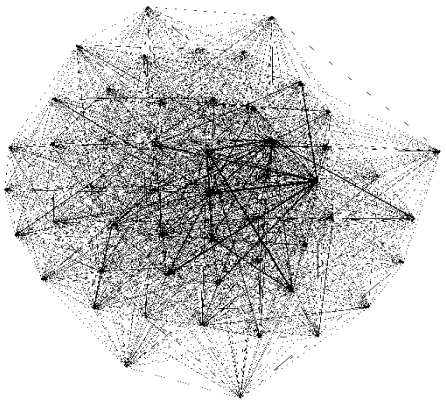


3.1. OpenOrd: In this layout, the nodes were placed on the lattice and nodes with a high degree in the center. But, technically this layout is a good fit for an undirected weighted graph to distinguish the cluster. As our network is directed, we will not use this layout.

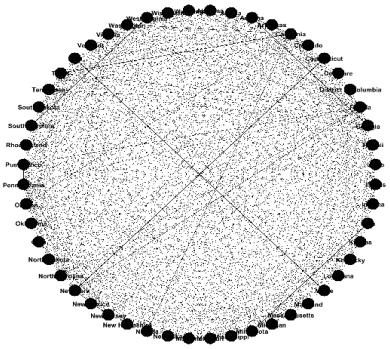


3.2. Fruchterman Reingold: In this layout, the nodes with less degree were placed on the perimeter, and nodes with a higher degree were placed towards the center of the sphere. It is denser than Yifun Hu in the center. Hence, we will not be using it in this analysis.

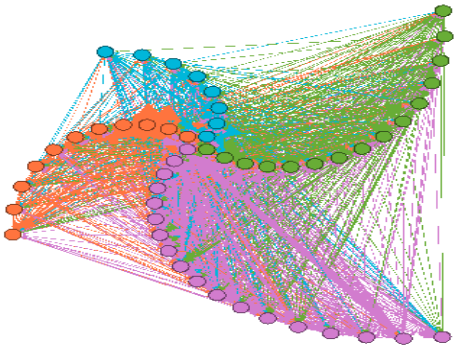
3.3. Yifan Hu: In this layout, the nodes with a higher degree are placed in the center and the nodes with lower degrees and fewer weights are pushed away from the center. It creates a repulsion based on Barnes-Hut's algorithm, where it calculates a repulsion of a node from a cluster (treat it as a single node) very effectively. It runs fast and stops automatically we will use this for our analysis.



3.4. The circular layout: This layout is very important to rank the nodes based on degree, weights, and other statistics. We can arrange the nodes on the circle: clockwise, anticlockwise.



3.5. The Radial Axis Layout: This layout places the nodes on the radial axis and with the least possible edge overlapping. Also, it can be grouped based on statistics such as degree, in-out-degrees, modularity, etc. We used this layout to find the communities and homophily in the state to state migration network.



In this paper, we used three layouts based on the requirement of the analysis. But for the beginning of the analysis, to study the migration patterns and to visualize the contributor or recipient of the migration, we have used the Yifan Hu layout. Then, the circular layout was used to rank the states and metropolitan areas, based on their net migration, in-migration, or out-migration. Lastly, we have used the radial axis layout to visualize the homophily and communities in the network.

4. **Sizing, Coloring, and Naming the nodes**: We have sized the nodes based on various statistics as discussed below and colored the nodes based on the nature of extracted information. Furthermore, we have colored and ranked the edges based on their weights which helped us to see the in-migration and outmigration using the bold and big arrow marks.

In figure 4.1, we have sized the nodes based on the gross migration (in-migration + out-migration) and ranked the edges based on the weights which enable the bigger arrows of out and in-degrees. This graph provides good information about the states with high mobility in the United States, the bigger the node, the more is the mobility. For instance, California, Florida, Texas, New York, and North Carolina are among the states with the highest mobility.

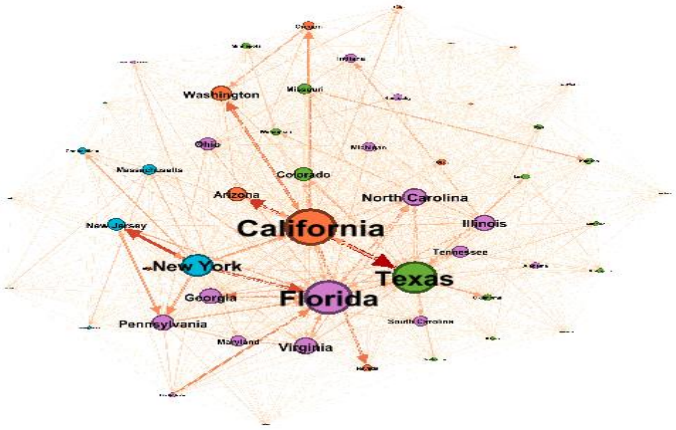


Figure 4.1

We have obtained many valuable pieces of information and direction for further analysis from sizing, coloring, and naming the nodes as discussed below in section 5:

5. **Relevant Statistics**: As we have discussed in section 2 that we have run all the statistics, but the centrality measures like betweenness, closeness, and eigenvector centralities will not be effective measures as First, migration from one state to another state is independent of other migration. Second, the degree of each node is almost the same so eigen centralities will not be an effective measure.

   However, the most useful statistics for this analysis are: first, degree statistics such as in-degree, out-degree, weighted in-degree, weighted out-degree, net weighted in-degree, net weighted out-degree, gross in-degree, and gross out-degree. And, the second useful statistic is the modularity class in detecting the communities and homophiles.

In the following analysis, we have tried to find the hubs in the states and hubs in the metro areas of the states based on the high inflow or outflow in the state. Also, we have tried to find the top contributor to the state's in-migration, and recipients of the state's out-migration. Let's see the analysis as follows:

1. Hubs based on inflows in the state to state migration and top contributors
2. Hubs based on outflows in the state to state migration
3. Hubs in the top 5 states with the highest inflows(in-migration) in metro to metro migration
4. Hubs in the top 5 states with the highest outflows(out-migration) in metro to metro migration

5.1.Hubs based on in-degree: In fig 5.1.1, we have plotted the Yifan Hu layout and sized the nodes based on the in-degree migration and to rank them we have used the circular layout in figure 5.1.2

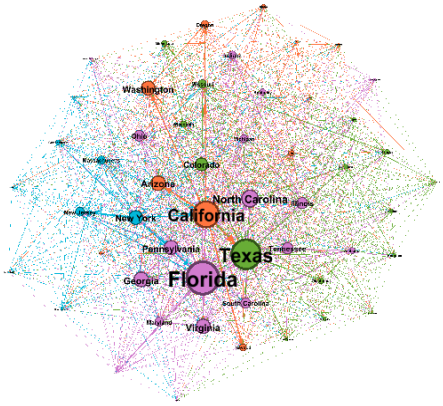
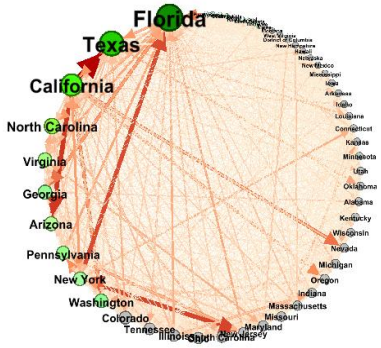


Figure 5.1.1



Label	FL_contri
Florida	New York
Texas	Puerto Rico
California	Georgia
North Carolina	Virginia
Virginia	
Georgia	TX_contri
Arizona	California
Pennsylvania	Florida
New York	Louisiana
Washington	

Figure 5.1.2

5.2.Hubs based on out-degree: In fig 5.2.1, we have plotted the Yifan Hu layout and sized the nodes based on the out-degree migration and to rank them we have used the circular layout in figure 5.2.2

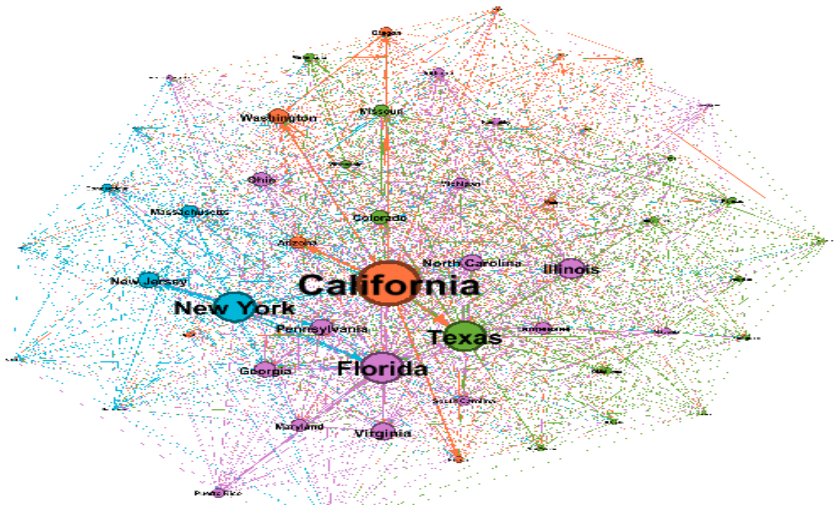


Figure 5.2.1

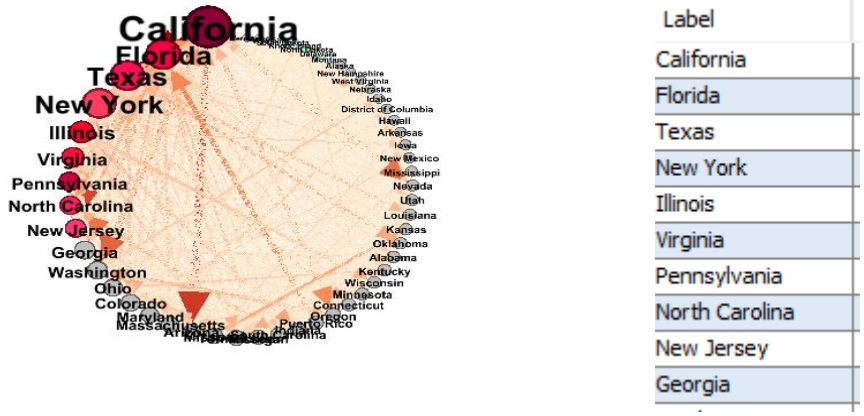


Figure 5.2.2

5.3.Hubs based on the net out-degree: In Figure 5.3.1 and Figure 5.3.2, we have plotted the Yifan Hu layout and sized the nodes based on the net out-degree migration and net in-degree migration, where net migration = in-degree – out-degree.

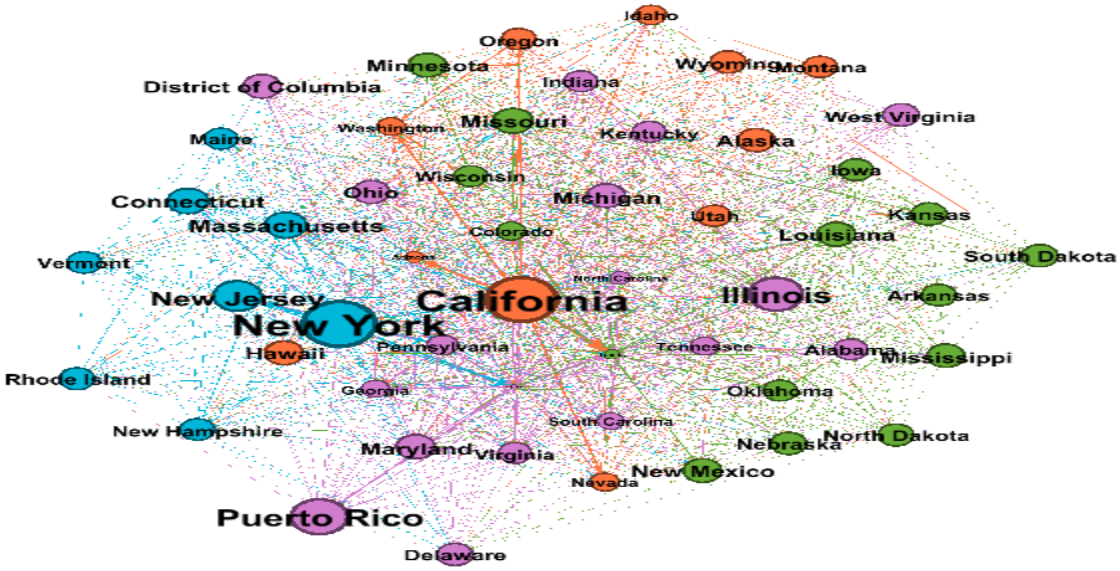


Figure 5.3.1

In the following figure, purple nodes are the top 5 states in the net in-migration.

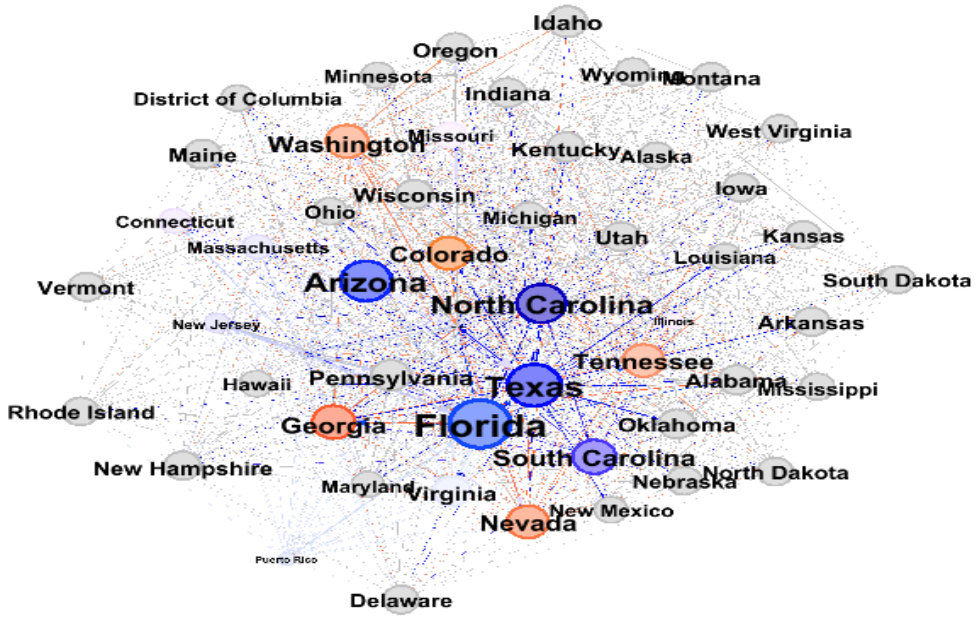
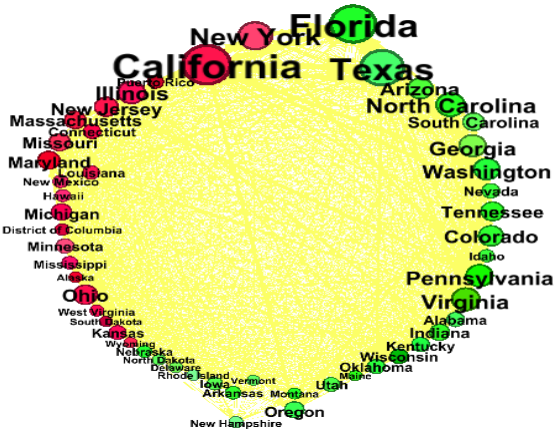


Figure 5.3.2

In Figure 5.3.3, we used the circular layout to rank the nodes based on in-degree migration, clockwise. Also, we have sized them based on in-migration (Figure 5.1.1, Figure 5.2.2) and colored state nodes red if net migration is negative and green if net migration is positive. We sized them according to in-degree migration as net degree migration for California and New York will be the lowest and make them very tine nodes in the circular graph.

Top 10 Sates with net outmigration

Label
New York
California
Puerto Rico
Illinois
New Jersey
Massachusetts
Connecticut
Missouri
Maryland
Louisiana



Top 10 Sates with net inmigration

Label
Florida
Texas
Arizona
North Carolina
South Carolina
Georgia
Washington
Nevada
Tennessee
Colorado

Figure 5.3.3

The following table shows the biggest recipient of out-migration in the states: New York, California, Puerto Rico, and Illinois. This will answer the question, if people are moving out of the state then where are they heading?

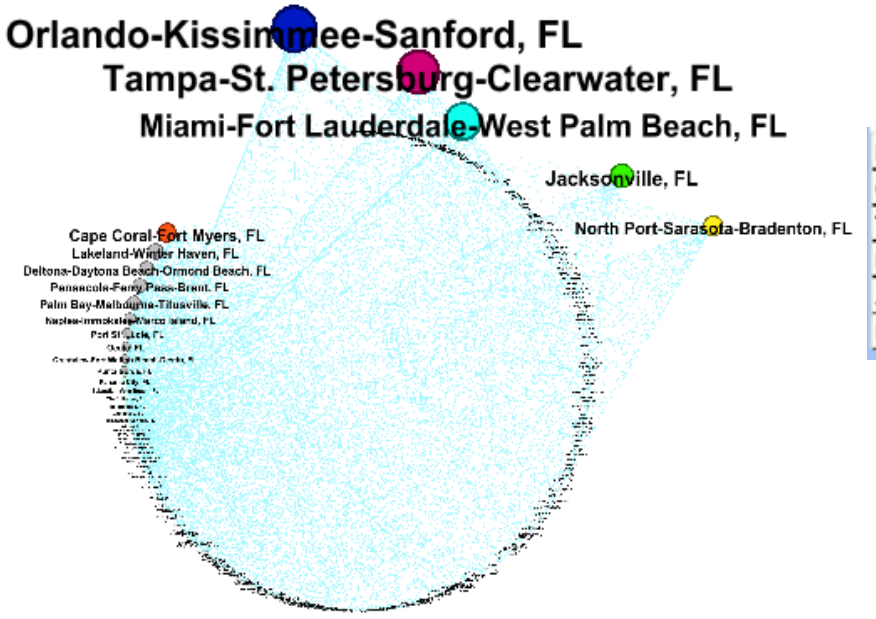
target_NY	target_CA	target_PR	target_IL
New Jersey	Texas	Florida	Indiana
Florida	Arizona	Pennsylvania	Florida
Pennsylvania	Washington	Texas	Wisconsin
	Nevada		

5.4.Metro to Metro migration analysis: The following graphs show the metropolitan areas in the particular state gaining population from other states. In this study, we discarded the metro to metro migration within the state. Also, we sized, colored, labeled, and ranked the metropolitan areas for the top 5 states with the highest in-migration and highest out-migration.

a) Metro to Metro in-migration:

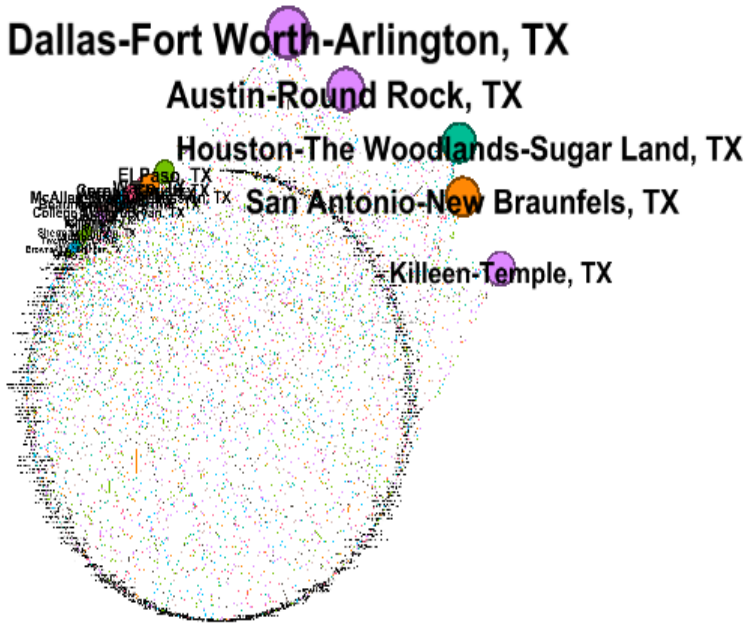
Top 5 states with the highest inflow and their contributors: After studying state to state migration and analyzing top states. Now, we would like to know that if people are moving to these states then in each of these top states which metropolitan areas (moving in hubs) are gaining the most?

1. **Florida:**



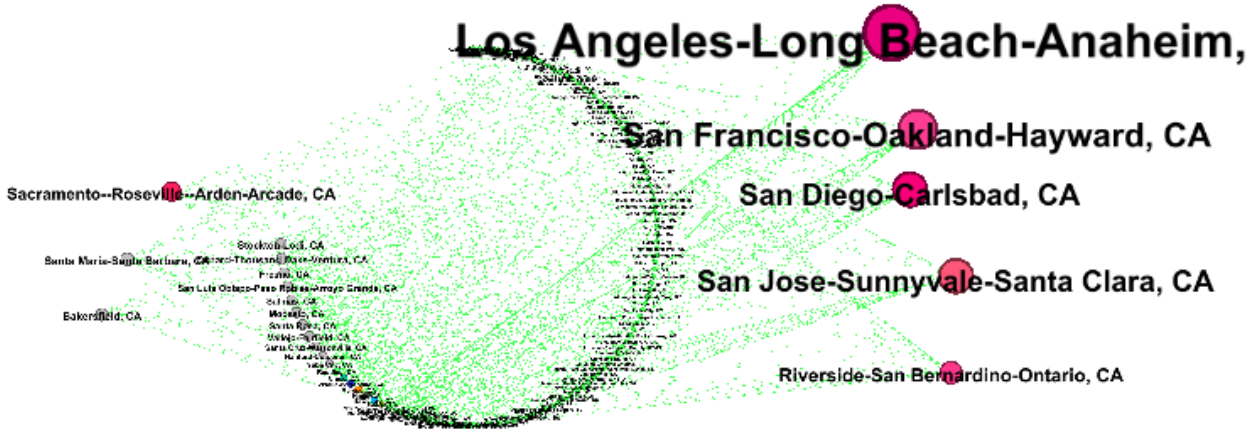
Florida_hubs
Orlando-Kissimmee-Sanford, FL
Tampa-St. Petersburg-Clearwater, FL
Miami-Fort Lauderdale-West Palm Beach, FL
Jacksonville, FL
North Port-Sarasota-Bradenton, FL

2. **Texas:**



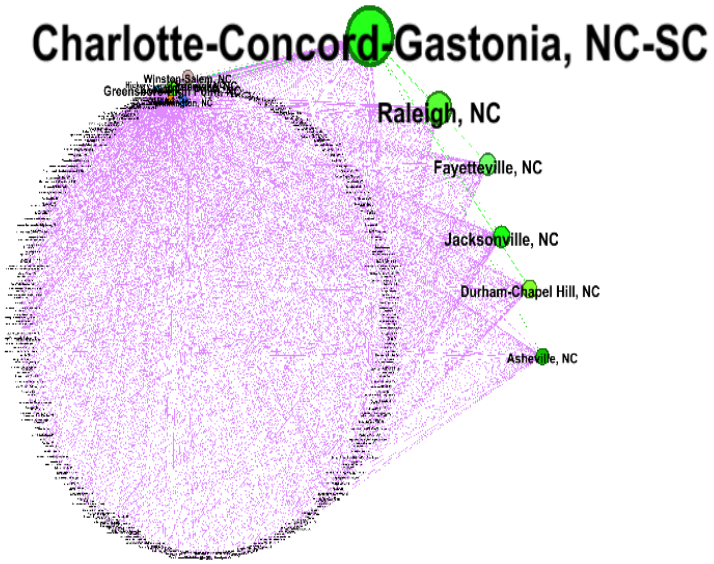
Tx_metro_hub
Dallas-Fort Worth-Arlington
Austin-Round Rock
Houston-The Woodlands-Sugar Land
San Antonio-New Braunfels
Killeen-Temple

3. **California:**



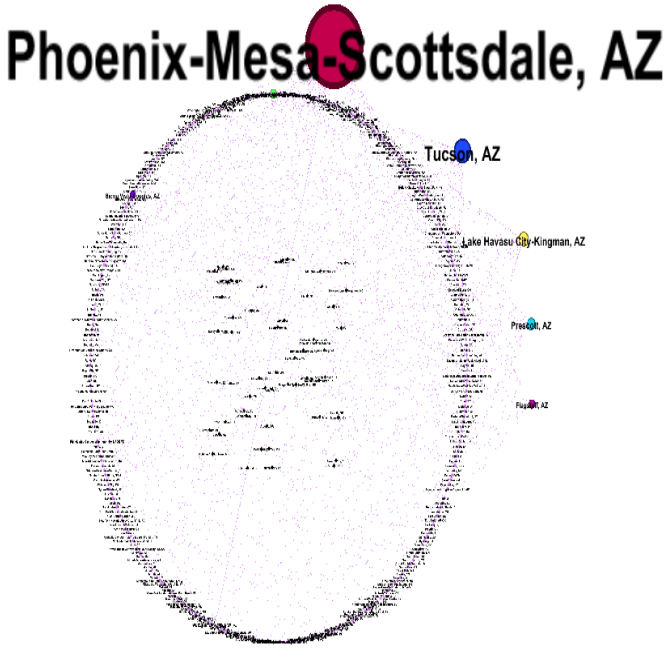
CA_hubs
Los Angeles-Long Beach-Anaheim, CA
San Francisco-Oakland-Hayward, CA
San Diego-Carlsbad, CA
San Jose-Sunnyvale-Santa Clara, CA
Riverside-San Bernardino-Ontario, CA

4. **North Carolina:**



NC_hubs
Charlotte-Concord-Gastonia, NC-SC
Raleigh, NC
Fayetteville, NC
Jacksonville, NC
Durham-Chapel Hill, NC

5. **Arizona:**

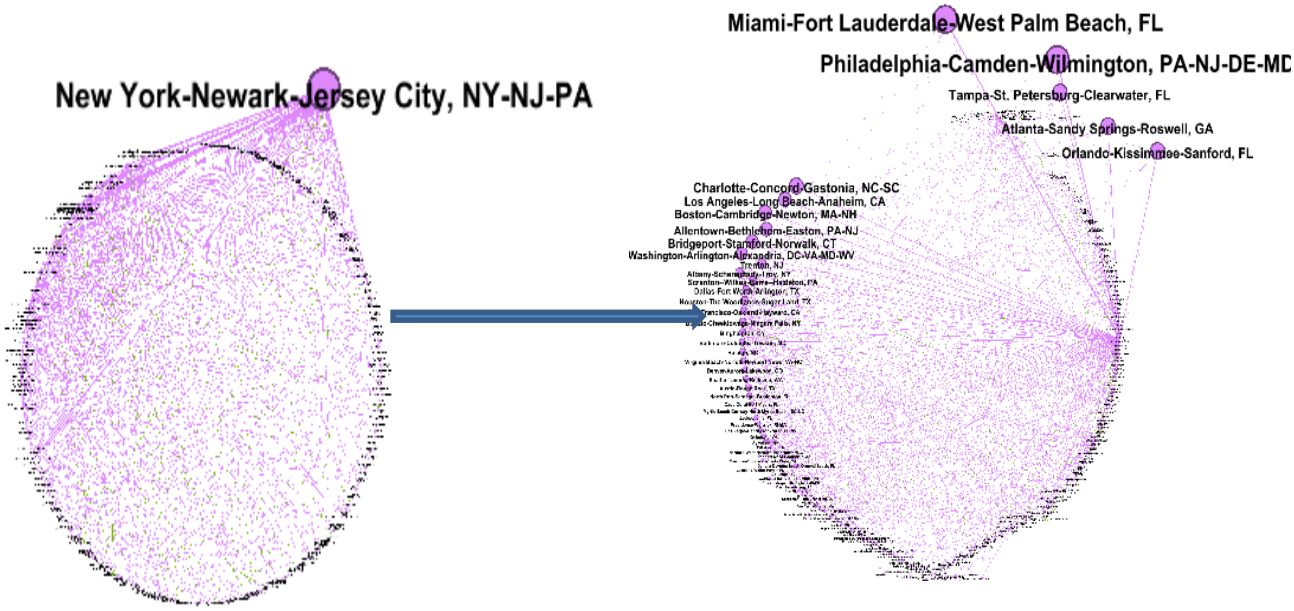


Arizona_hubs
Phoenix-Mesa-Scottsdale, AZ
Tucson, AZ
Lake Havasu City-Kingman, AZ
Prescott, AZ
Flagstaff, AZ

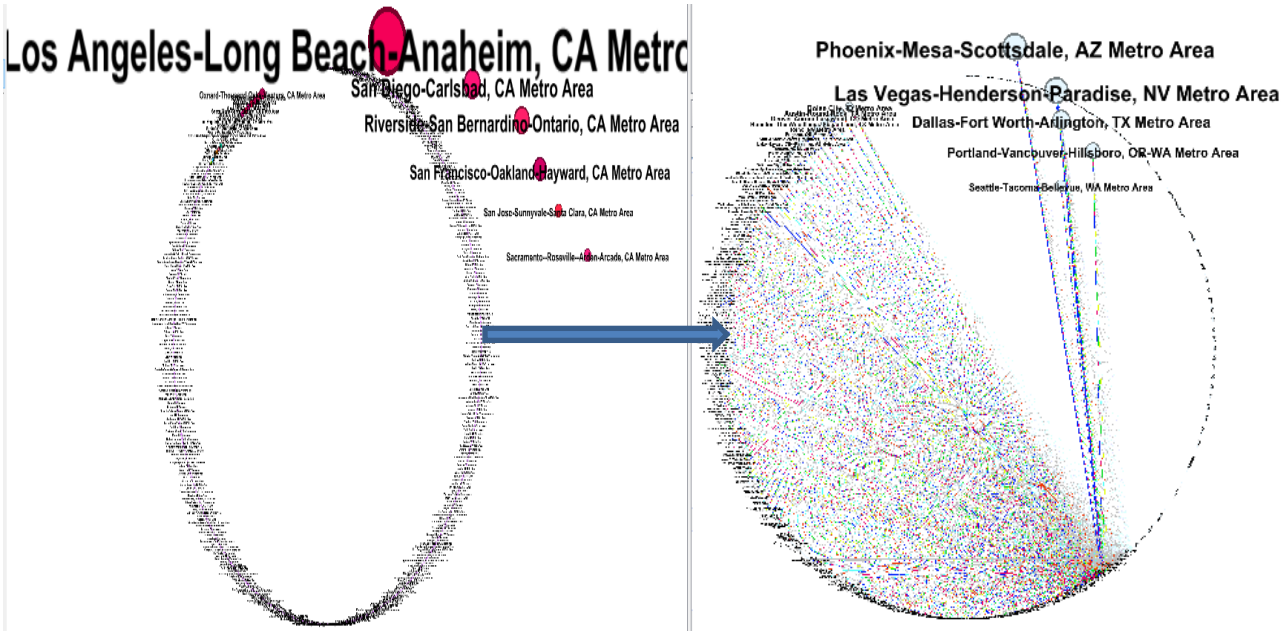
b) **Metro to metro out-migration:**

Top 5 states with highest out-migration and their recipients: As we have looked at the biggest contributors to state's in-migration, we studied the top states with highest out-migration and if people are moving out of these states then which metropolitan area do they prefer?

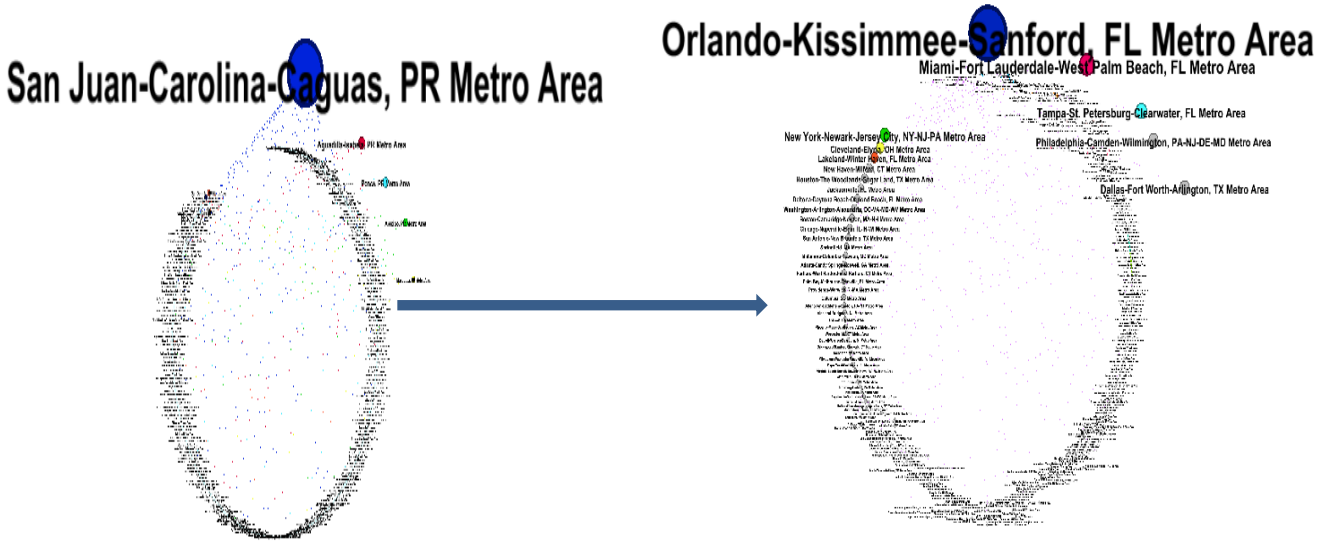
1. New York:



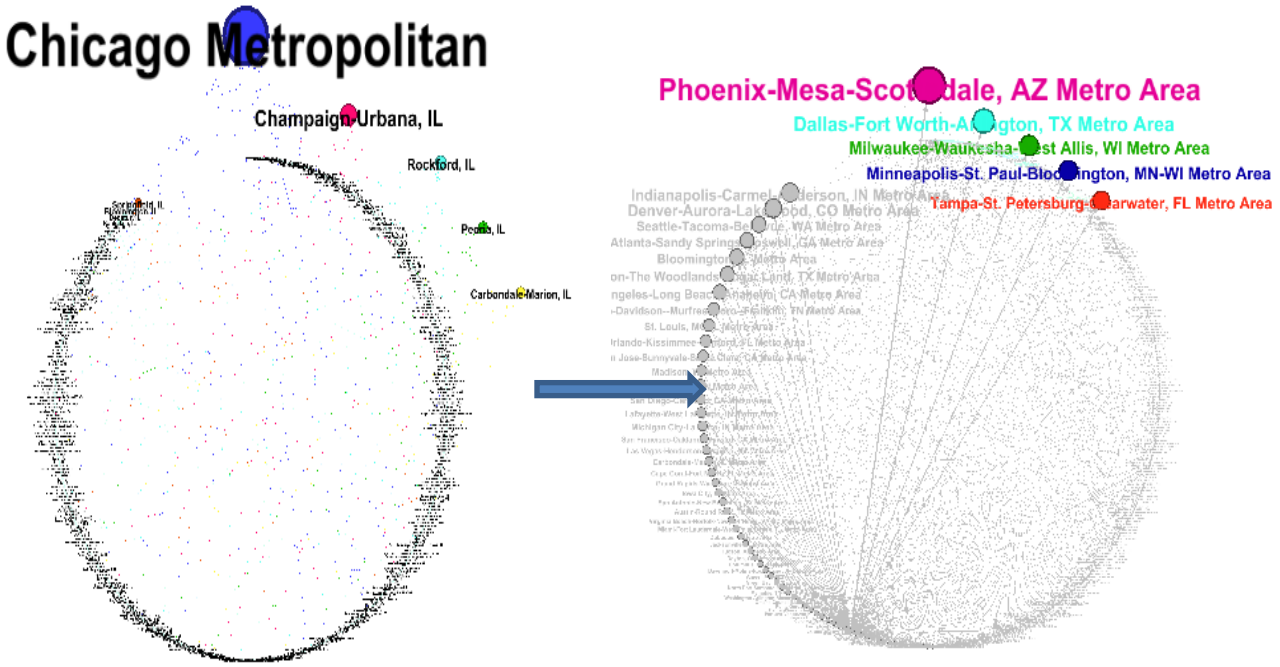
2. California



3. Puerto Rico:



4. Illinois:



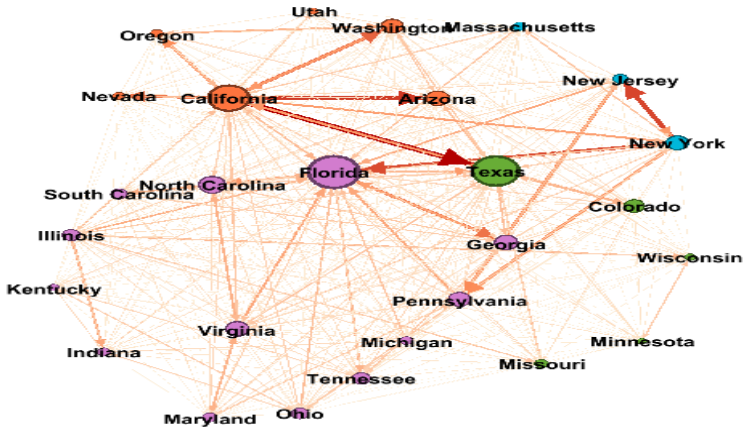
5. **New Jersey:**

Based on statistics, we found the top 10 states with high mobility and the top 10 states with the lowest mobility. Furthermore, we were able to rank all the states based on their total net migration (net weighted degree of a node) in the year 2018. Also, using the data laboratory and graph visualization we listed top contributors to the states in-migration and top recipients of the stats' out-migration.

Another important statistic was modularity; we were able to see the pattern of the regionwise migration and same coastline migration in the United States. We will discuss more details in the following sections 6 and 7.

6. **Filters**: As this study is mainly focused on high mobility states, we need to filter the states with low mobility. We used a degree filter (gross degree = 95) to narrow down our study to high mobility states and to form another network of the metro to metro migration in those selected states. Also, we applied a modularity filter on the degree filter to understand the migration patterns as shown in Figure 6.1. As shown in the figure, we have only 27 states and 4 communities among these states.

a) Degree filter and a graph:



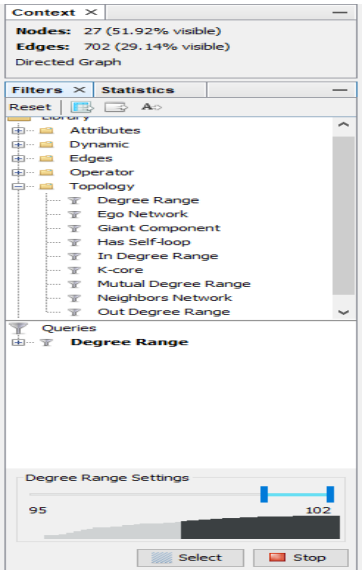
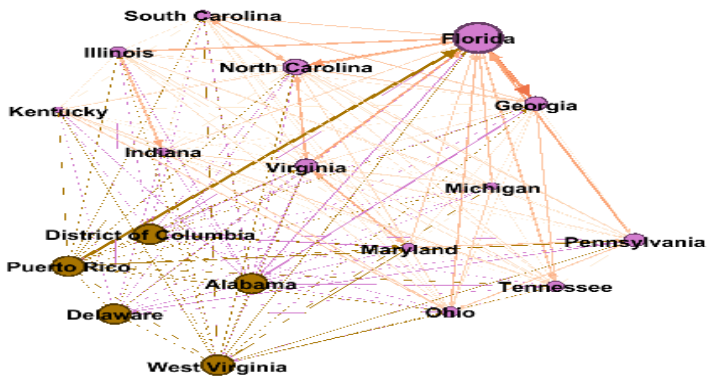
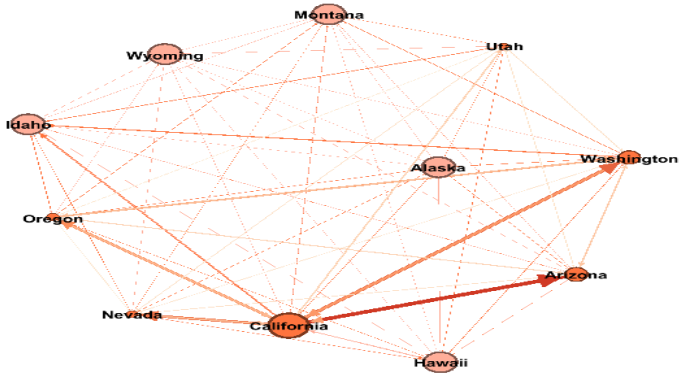


Figure 6.1

b) Modularity on degree filter: Each of the following graphs in the table represents the communities in the migration network.

This graph shows one community. If we observe these states on the US map then they lie at the same longitude. More precisely they are eastern central and south Atlantic states as shown in Figure 7.2.	
In this community, all the states are from the west region. Therefore, this community is homophily.	

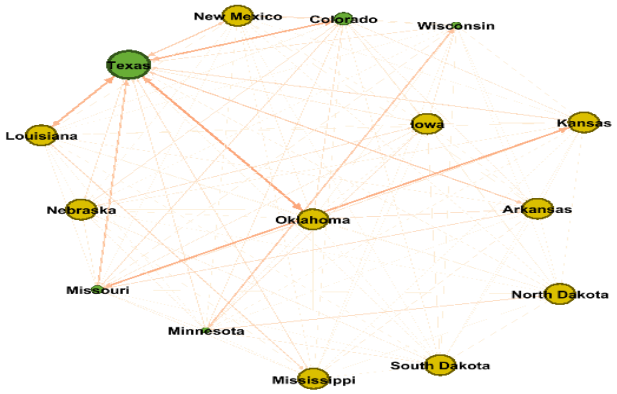
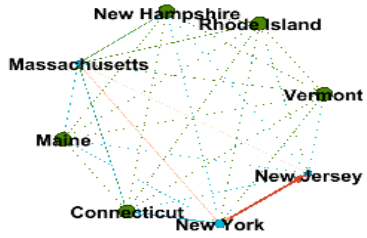
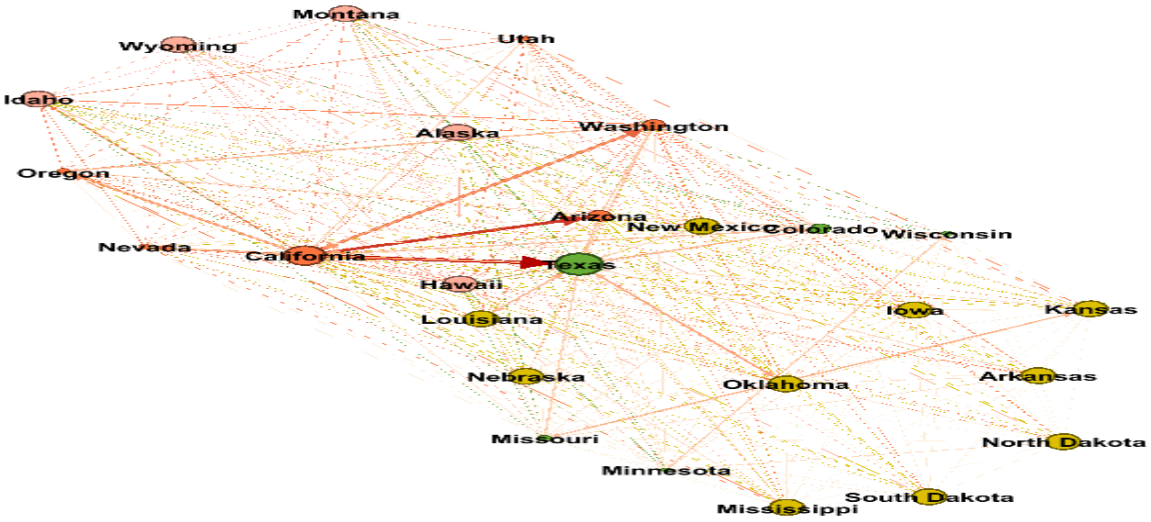
In this community, all the states are from the west north-central and the west south-central regions. Also, they are located around the same longitude too.	
This community consists of all northeastern states. Hence, there is homophily too.	

Table 6.1

c) Lastly, if we filter 2 modularity classes of the western region. Then it is very clear and confirms our finding that western states prefer to migrate to western states. As shown in the following figure, a combination of two communities indicates that there is high mobility in the west, west north-central, and west south-central regions.



7. **Features: communities, giant clusters, homophily, robustness, spreading phenomenon**:
In this study, the most useful feature was communities and homophilies which were used to discover the migration patterns. Using the modularity statistics and coloring the nodes using modularity, we could find the regional migration pattern as shown in the following Figure 7.3. That suggests that people tend to live in the same region or move to the neighboring states. As discussed in the above table with modularity filters and shown below, there are 4 communities in the United States. Also, there are regionwise homophilies (e.g. west and northeast) and longitude wise homophily (e.g. east-central and west-central regions). From figure 7.3 (shown in dark brown color), There is a homophily in east coast states (e.g. South Atlantic and Middle Atlantic regions).

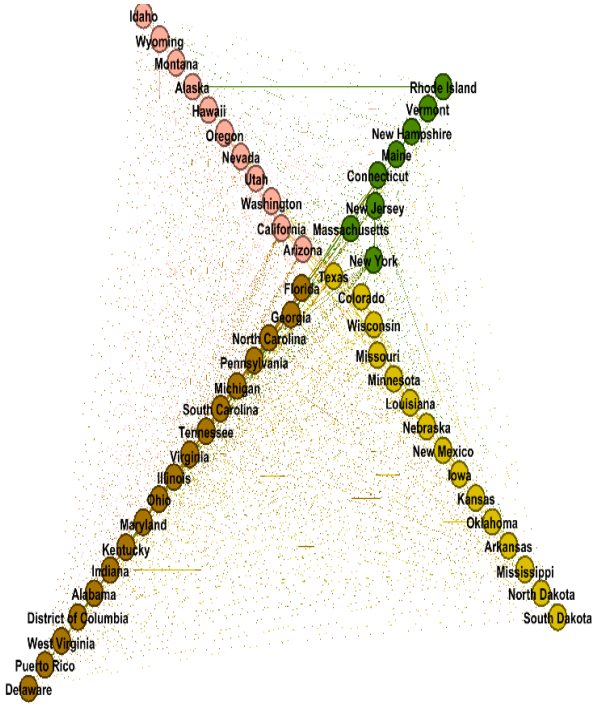


Figure 7.3

	Northeast	Midwest	South	West
2018-2019	-238	17	263	-41
2017-2018	-352	-38	512	-122
2016-2017	-200	-25	44	181
2015-2016	-216	-71	39	247
2014-2015	-101	-112	271	-57

Figure 7.1

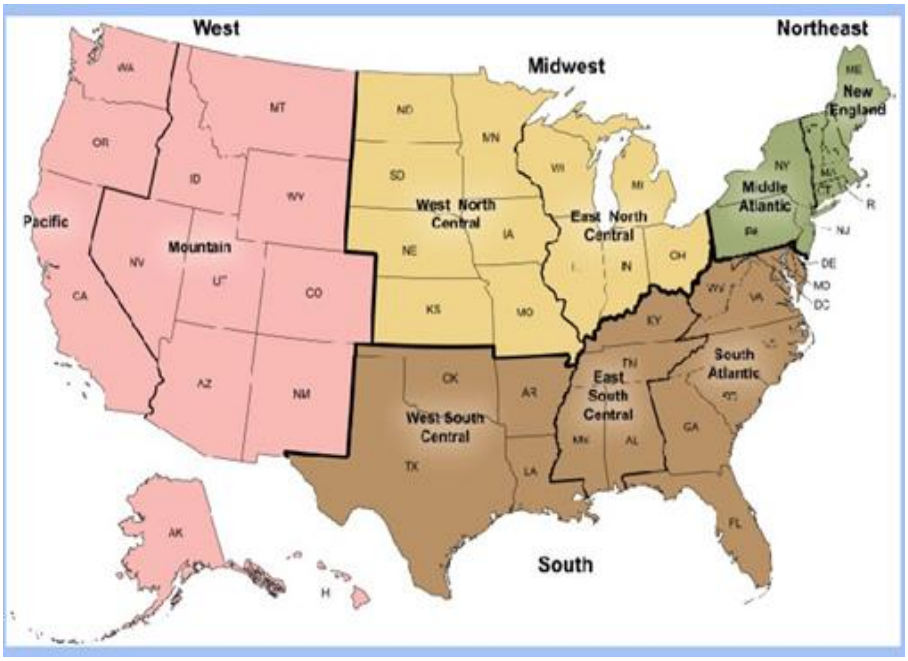


Figure 7.2

8. Result and contemplation:

8.1. Results: Based on the study of the regional migration, there are more outflows from the northeastern region and inflow toward the southern region. Furthermore, we have analyzed and ranked which states have more net inflow and which states have more net outflow. Moreover, we found and ranked hubs in these high inflow states and out moving hubs in the high outflow states as shown below:

Highest Net Migration

Label	Florida_hubs
	Orlando-Kissimmee-Sanford, FL
	Tampa-St. Petersburg-Clearwater, FL
	Miami-Fort Lauderdale-West Palm Beach, FL
Florida	Jacksonville, FL
	North Port-Sarasota-Bradenton, FL
Texas	Tx_metro_hub
	Dallas-Fort Worth-Arlington
Arizona	Austin-Round Rock
	Houston-The Woodlands-Sugar Land
North Carolina	San Antonio-New Braunfels
	Killeen-Temple
South Carolina	Arizona_hubs
	Phoenix-Mesa-Scottsdale, AZ
Georgia	Tucson, AZ
	Lake Havasu City-Kingman, AZ
Washington	Prescott, AZ
	Flagstaff, AZ
Nevada	NC_hubs
	Charlotte-Concord-Gastonia, NC-SC
Tennessee	Raleigh, NC
	Fayetteville, NC
Colorado	Jacksonville, NC
	Durham-Chapel Hill, NC

Figure 8.1

Lowest Net Migration

Label	Out-MovingHubs
New York	New York-Newark-Jersey City, NY-NJ-PA
California	Los Angeles-Long Beach-Anaheim, CA
Puerto Rico	San Juan-Carolina-Caguas, PR
Illinois	Chicago-Naperville-Elgin, IL-IN-WI
	Trenton, NJ
New Jersey	
Massachusetts	
Connecticut	
Missouri	
Maryland	
Louisiana	

Figure 8.2

In this analysis we have attempted to answer three questions:

1. **What are the domestic migration patterns in the United States?**
   Based on US census data, we found a homophily in the regional migration as shown in the above Figure 7.3 and Table 6.1. Also, using filters we found that northeastern states have more neighboring state migration, for example, New York, Pennsylvania, New Jersey, and Massachusetts. Also, western states prefer to migrate to western neighboring states, for example, California, Arizona, Washington, Oregon, and farthest central west Texas. Furthermore, people in the east coast states, including northern and southern states, prefer to stay on the east coast of the United States.

2. **Where people are moving?**

As shown in Figure 7.1 and Figure 8.1, People are moving to southern states, for example, Florida, Texas, North Carolina, South Carolina, Georgia are among the top 10 states with the highest net migration. Also, some of the southern states are appealing to more people too. For example, Arizona, Washington, Nevada, and Colorado are among the top 10 states with the highest in-migration as shown in Figure 8.1.

On the other hand, data suggests that people are moving out of the northeastern states like New York, New Jersey. Even though western states are gaining, but net migration in some of the states is getting lower and lower within the last 5 years for example California; California is among the second largest states with high out-migration. According to United Van Lines, Illinois is on the list of the states with the highest out-migration since 2010. It is ranked 4th in our list of states with low net migration.

Lastly, Top destinations in the migration pattern are Central Florida (Orlando, Tampa) and Texas (Dallas, Austin). Most losing metropolitan areas in the United States are New York and Los Angeles.

What are the causes and effects of domestic migration in the US?

1. **Causes of domestic migration:**
   a) **Causes for overall migration**: We have studied the dataset of region wise migration reasons and overall US internal migration reasons for the years (2014-2018). They have divided the migration of 32 million people annually into 4 categories: family-related, job related, housing-related, and others. Out of these 32 million migrating people, 30% is family-based (married, divorce, wanted to stay close to the family, etc.), 20% is job-related migration(new Job, first Job, retirement, job transfer), 45% of the migration was based on housing-related (grater space, good neighborhood, low-cost living, wants to own the house, etc.) and the last category of others (climate change, warmer weather, natural disaster, etc.).
   b) **Causes for Interstate migration**: As we have studied the state to state migration, we wanted to know the interstate reasons and from the above data we have extracted data for different state migration as follows: First, every year 3.5 million people annually migrate statewide due to job-related reasons (a new job, a first job, job transfer). Second, 2 million people migrate to stay closer to work or easier commute to workplaces. Third, 1 million people migrate statewide to attend or leave college. However, these numbers will change for the year 2020, due to the current pandemic. Fourth, every year 350,000 people migrate to different states due to retirement (moving to the home state or warmer states in the southern region). Lastly, on average 400,000 people migrate to different states due to climate change and natural disasters, this number has doubled in the last

two years and researches are suggesting that it will get worst worldwide in this century [3].

In conclusion, In the United States people move from one place to another due to many reasons as mentioned above, but the next year's statistics will give us skewed or accelerated results due to Covid-19.

2. **Effects of domestic migration:** There are a couple of obvious effects of the migrations like infrastructure, real estate, and labor market. For instance, research shows that greater space and lower cost of living in the states like Florida, Oregon, the Carolinas, Washington, Texas, Arizona, and Nevada are in the center of migration and are the causes of the real estate boom in theses states [1]. Furthermore, domestic in-migration helps the state to lower the local unemployment rate substantially for many years. Also, due to high in-migration to swing states like Florida, Nevada, North Carolina, and Arizona will become more important in deciding the administration for the White House due to increasing electoral votes. On the other hand, real estate sites like Zillow and Realtor.com are showing the high inventory of the houses in the northeastern states and California. In particular, we will see more empty houses in metropolitan areas like New York and the LA metropolitan areas. [2].

   Also, from the above graph, we can see that Millenials are move to southern and westerns regions which will impact the labor and housing market in those regions. Also, Generation x is flocking to the southern regions to retire in warmer places which will impact the healthcare industry.

Region	North East	south	west	mid west
age groups	different states	different state movers	different state movers	different state movers
20 to 24 years	113	228	180	127
25 to 29 years	157	269	189	188
30 to 34 years	77	210	119	111
50 to 54 years	35	105	66	21
55 to 59 years	40	81	45	59
Total 1 year	686	1,831	1,213	1,008

8.2.Final contemplation: We have learned a lot in this network science class and using the knowledge taught in the class, we were able to find results same as a published article for the company's like United Van Lines, Zillow.com, and articles published on Census.gov. This gives a kind of satisfaction to the output of this analysis.

Additional Data: Although we have used data from the Census.gov, some of the data for the reasons were incomplete or were not done. For instance, they didn't study the state to state migration reasons explicitly which could have enhanced the study. For example, statewise migration based on age, sex, education, climate change. Also, as we mentioned earlier that we wanted to do a migration study based on Covid-19, but due to a lack of data, we couldn't work in that direction. Additionally, this is a pandemic year and we

have faced many challenges and changed our lifestyle completely 180 degrees, 2020 migration data and reasons will be very interesting to study. Furthermore, This year 2020 is also special in the sense of census count too, new and current population count will make more precise results of the migration with less margin of errors.

Data preparedness: We got raw data from census.gov, first we had to arrange our thoughts and look for the relevant data to study. We did a lot of research about the migration patterns within the United States and try to find the most current data from other sources, but that would be a subset of the census.gov. Therefore, we decided to go with the census data. After selecting the data, the next step was to convert the data into excel/.csv format as a data feed to Gephi. We removed the irrelevant data and made edge and node input files for Gephi. Later, we realized that to find the hub in the state we need more data, and decided to study another network of the metro to metro migration in the United States and repeated the process of data cleaning and preparing the input file for Gephi.

Unexpected: State to state migration data file turns out to be a matrix format, so when we import it in Gephi, we got wrong results as it considered the target node as the source node. Also, later in the analysis, we needed net migration of a state, which was not in the data and Gephi doesn't calculate it. So we exported all the calculated statistics in Excel, using weighted in and weighted out-degree, we calculated the net weighted degree of a node and then imported that file in Gephi with some variable type adjustments. Also, importing the whole metro to metro dataset was not possible due to memory issues, Then we studied each state's migration network individually by filtering in-migration states/out-migration states. Finally, to understand the network very well, we had to convert our undirected network to a directed network. We had to deal with negative weights in the metro to metro analysis.

Studying network science and applying learned topics to analyze the real data was a very satisfactory outcome of the course. Also, we look forward to learning more about real data analysis using network science techniques.

Data source:

a) Region migration: https://www.census.gov/data/tables/time-series/demo/geographic-mobility/historic.html

b) State to state migration: https://www.census.gov/data/tables/timeseries/demo/geographic-mobility/state-to-state-migration.html

c) https://www.census.gov/topics/population/migration/guidance/metro-to-metro-migration-flows.html

d) Reasontable:https://www.census.gov/data/tables/time-series/demo/geographic-mobility/historic.html

e) Regionwise migration reasons: https://www.census.gov/data/tables/2019/demo/geographic-mobility/cps-2019.html

f) Moving companies: https://www.unitedvanlines.com/ ,

https://www.uhaul.com/

g) Real estate data: https://www.zillow.com/research/data/ ,

https://www.realtor.com/research/data/

References:

1. https://www.move.org/why-americans-are-moving-to-southern-and-western-states/
2. https://www.census.gov/library/stories/2019/04/moves-from-south-west-dominate-recent-migration-flows.html
3. Stephan A. Schwartz (2020, August), America, Covid-19, Climate change, and Migration, Explore
4. https://www.jchs.harvard.edu/blog/not-just-the-sunbelt-millennials-and-baby-boomers-increasingly-head-west
5. Cherven Ken, PACKT publishing(2015), Mastering Gephi Network Visualization, 978-1-78398-734-4
6. https://www.census.gov/topics/population/migration/guidance/metro-to-metro-migration-flows.html
7. https://www.census.gov/topics/population/migration/guidance/state-to-state-migration-flows.html
8. https://www.census.gov/library/publications/2014/demo/p20-574.html