Mexican Drug Traffic
Social Network Analysis

Even Thullen, Mike Hu, Yongjia Sun

Claremont Graduate University

May 16th, 2021

1. Introduction of the Mexico drug trafficking and Background literature

Mexico is the major player of illegal drugs for more than a century. Over the last few decades, drug production in Mexico has been gradually increased and Mexico now has become one of the strongest suppliers of illicit narcotics to the world 1. This high risk, high reward illegal business in Mexico has largely been controlled by only a few cartels. In addition, these few players in the game have created barriers to stay on top of the entire chain of production [1]. Sinaloa Cartel, for example, is one of the largest drug-trafficking organizations in the world. Sinaloa was founded in the late 1980s in north-west Mexico by El Chapo, also widely known as "Shorty". Sinaloa and "Shorty" have a reputation for "business first", meaning that they will get rid of anything stands in their way. Therefore, Sinaloa also had access to a huge variety of weapons, which had lad to countless violence activities [2]. A few key players like Sinaloa had rich resources, in terms of transportation. Roads, highways, trucks, maritime ports, and airports are frequently used to smuggle drugs [1].

Things started to change when Cocaine and synthetic drugs entered Mexico. This has led to the rise of new gangs, which have begun battling traditional Mexican cartels for a share of the smuggling business [1]. The war between drug cartels is serious. To create reputation and barriers for new players, old cartels constantly kidnapped, tortured, and murdered members of rival criminal organizations [2]. At the same time, old cartels had quickly joined the business of cocaine and synthetic drugs because of the huge profit margin. New competitors and existing cartels soon emerged and redefined their routes for drug activities [1].

The emerged and redefined drug trafficking route is crucial to our study. After years of battling with rival drug gangs and law enforcers, drug smuggling routes are chosen to maximize profit, while minimizing costs. In other words, drug trafficking chooses their business operation in a very rational manner [1]. Drug transportation costs is only one aspect, there are some other important factors. It is worth mention here that risk of a particular route can turn into significant costs. With that being said, there must be patterns and associations among all the drug smuggling routes, which lead to our research question:

What is the association and patterns, between Mexican municipalities and drug traffic organization?

We hope our findings can contribute and support the understanding of the drug trafficking routes in Mexico. Moreover, our project serves as a demonstration of what we have learned throughout the course. Using R as our main system, we have built an inexpensive, easy to replicate, but also provides great details for interpretation.

2. Background literature

Here are two main literatures we found that are related to the focus of our study.

Mexico's drug networks: Modeling the smuggling routes towards the United States, Applied Geography (2014)

- According to the study of Mexico's drug networks, there are many different drug shipment methods [1].

- Small amounts that can be transported by car, SUV or pickup truck, to loads of one ton or more that are packed into commercial and non-commercial vehicles [1].
- criminals do not make random movements, neither do crimes occur randomly at places. Smugglers are believed to have a certain set of preferred routes for moving the different types of drugs [1].
- On the one hand, the smugglers are actively considering and evaluating new opportunities for smuggling routes. they are preys who try to sneak around avoiding being captured by police or killed by other cartels [1].
- The pattern of the drug trafficking routes is essential for effective allocation of police resources and for necessary implementation of regional policies and practices, in to order to crack down drug smuggling and drug violence [1].

Knowing Where and How Criminal Organizations Operate Using Web Content, Harvard Business (2012)

- In recent years, the power of telecommunication, transportation and technology has fostered an impressive growth rate in world complexity [3].
- Information complexity critically affects the ability of security agencies to collect intelligence information by making it more costly [3].
- The focus of this study is to bring the benefits of tracking complex phenomena in a inexpensive way [3].
- Exploiting some already indexed reliable sources such as online newspapers and blogs . (Note that although this article has claimed that their sources such as online news papers and blogs are reliable. We think otherwise and will be discuss later in the limitation section [3].

3. Relevant social network theory

For the relevant social network theory of our study, we are using affiliation network (exhibit 1). Affiliation network, also known as two mode networks. There are always two types of nodes in affiliation networks: one type for the actors, and the anther type for the groups or events to which the actors belong. Ties connect from the actors to the according groups. There are no direct ties among actors, and there are no direct ties between the groups. The data matrix for two-mode networks is an incidence matrix, and it depicts how n actors belong to g groups. For example, we have six students, which we would group them into four classes.

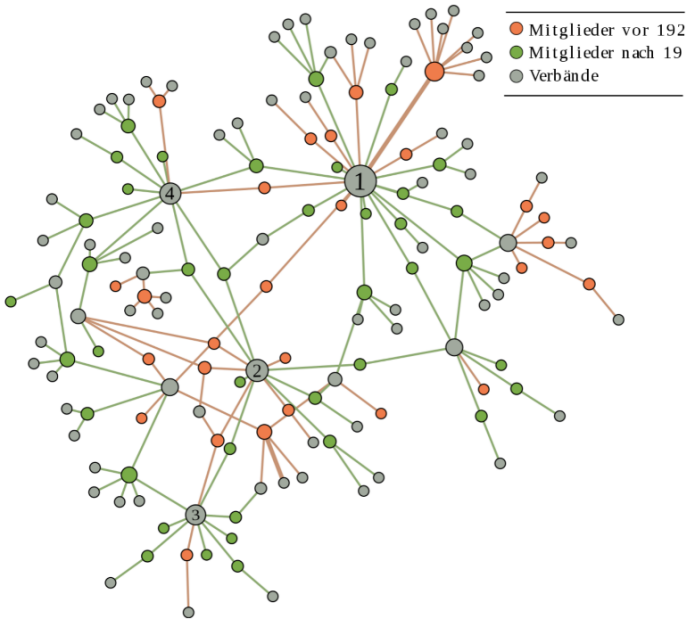


Exhibit 1 Affiliation Network

We can also use one-mode network in our study. One mode network has only one type of node. Those nodes have directed ties. The data matrix for one-mode networks is adjacency matrix. Adjacency matrix is an "nxn" square matrix where each dimension refers to the actors in the network. An incidence matrix, on the other hand, is an "nxg" rectangular matrix with two different dimensions: actors and groups.

4. Description of the data set

Our raw dataset is a temporal bipartite network capturing the association between Mexican municipalities and drug traffic organizations, represented as annual snapshots from 1990 – 2010, as extracted from Google News articles. This matrix CSV file contains a total number of 2,467 nodes and 6,228 edges. We felt comfortable with the size of our data set. This data set is large enough to provide significant findings and prevents our CPU for overload at the same time. There are a total of 4 attitudes, or variables:

Name	Description	Value	Data Type
Code (Column #1)	The zip code of the municipality in Mexico referred by the row.	This data set comes with coded zip that each zip code represents a city. There are a total 2457 different cities in our dataset.	Integer
State (column #2)	The INEGI code of the state in which the municipality is located.	This data set comes with coded number that each number represent a State. There are a total of 32 different state in our dataset.	Integer
Year (column #3)	The year to which the row refers. Note that data prior to 2004 is still reported, but less reliable than data after 2004	Data set of drug activities information are recorded from year 1990 to 2010.	Data/time
Drug group (Column #4-13)	A column per DTO (Drug Trafficking Organization). The column name identifies the organization. We collect the 9 largest and most important organizations, and we group in the last columns all mentions of the other DTOs.	There are a total of 10 drug cartels. The value in the data set "0" indicates no drug activities. The Value "1" indicates confirmed drug activities.	Integer

It is worth mention here that to better analysis the drug traffic, we had to split our original data set into 3 independent datasets. The original data set contains too many variables in one file that we found difficulty splitting using R. We decided to split them manually using excel, as shown in exhibit 3.

Exhibit 3 Data Split

	A	B	C	D	E	F	G	H	I	J	K
1	City	Beltran_Le	Beltran_Le	Familia	Golfo	Juarez	Sinaloa	Sinaloa_Fa	Tijuana	Zetas	Otros
2	1001	0	0	5	1	3	3	0	2	5	0
3	1002	0	0	1	0	3	2	0	0	4	0
4	1003	0	0	0	3	0	2	0	0	3	0
5	1004	0	0	0	0	0	0	0	0	0	0

	A	B	C	D	E	F	G	H	I	J	K
1	Year	Beltran_Le	Beltran_L	Familia	Juarez	Golfo	Sinaloa	Sinaloa_Fa	Tijuana	Zetas	Otros
2	1990	0	0	0	0	0	0	0	0	0	0
3	1991	0	0	0	0	3	3	0	0	0	0
4	1992	0	0	0	0	1	0	0	3	2	0
5	1993	0	0	0	2	1	12	0	6	3	0

	A	B	C	D	E	F	G	H	I	J	K
1	State	Beltran_Le	Beltran_L	Familia	Golfo	Juarez	Sinaloa	Sinaloa_Fa	Tijuana	Zetas	Otros
2	1	3	0	6	6	6	9	0	2	21	0
3	2	2	0	4	5	7	21	5	60	10	25
4	3	0	1	1	2	1	3	0	13	0	5
5	4	3	0	1	12	2	3	0	2	15	0

5. Visualizations of the chosen network and five-number summary

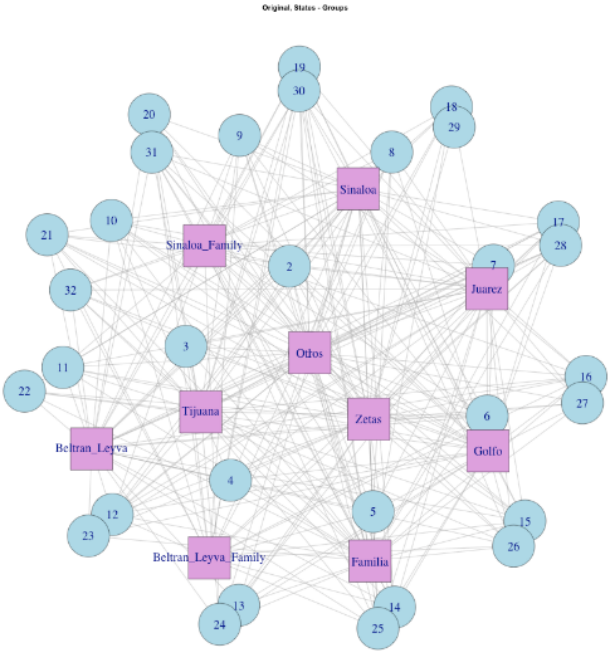
In this section, we will be presenting five number summary (node, edge, density, reciprocity, and components), degree/strength/weight analysis, visualizations, or plots, and examples on how we would use our results to understand Mexico drug activities. Network visualization, graph visualization and link analysis are the process of visually presenting networks of connected entities as links and nodes. Nodes represent the data points and links represent the connections between them. In this research project, we used R as the visualization tools hoping to better understand the connections among our variables. There are 2 alternative options that we have considered, which will be discuss in later section. As mentioned, we split the dataset into 3 sub-dataset bases on State, City, and Years. We want to gain a compressive knowledge about the drug activities using these three categories.

State

There are a total number of 42 nodes and 274 edges in the state dataset. The density, or the proportion of observed ties in number of possible ties is 0.32. The reciprocity is 1 and there are 1 component, meaning there is no subgroup. We then created an original network plot for state, as shown in Figure 4.1. The limitation of this plot is obvious. We can only see that each drug cartel has

some kind of drug trafficking operations, in many different states. A further analysis is needed here to get more insight. Therefore, we added degree and strength in our network for State.

Exhibit 4.1 Original Plot - State



More insights have revealed after we added degree and strength to our network. Table 1 and 2 from exhibit 4.2 represent the State degrees. According to table 1, drug cartel Sinaloa and Zetas both rank highest, which has 31 degrees following by cartel Beltram_Leyva, Familia, and Golfo. In other words, Sinaloa and Zetas had the widest network of doing drug activities across 31 different States. Table 2 shows that there are fifteen States that all ten of the Cartels have done drug activities. Table 3 and 4 are the strength matrix. Table 3 indicates that Zeta has the highest value of 1587 in strength, meaning that Zeta has done a total number of 1587 drug activities throughout all the States. We see that Zetas ranks the highest in both degree and strength. From table 4, State 16 ranks first, with a highest value of 541 in strength, meaning that there were 541 drug activities were happened in State 16. From our plot as shown in exhibit 4.3, we can easily see that Zetas has the largest square (indicates that Zetas ranks first in both degree and strength).

Exhibit 4.2 Degree and Strength Matrix - State

	group	degrees
6	Sinaloa	31
9	Zetas	31
1	Beltran_Leyva	30
3	Familia	30
4	Golfo	30
5	Juarez	28
8	Tijuana	27
10	Otros	24
7	Sinaloa_Family	23
2	Beltran_Leyva_Family	20

Table1

	state	degrees
2	5	10
3	6	10
4	8	10
5	9	10
6	10	10
8	12	10
9	14	10
10	15	10
11	16	10
12	17	10
13	19	10
14	21	10
15	25	10
17	28	10
18	30	10

Table2

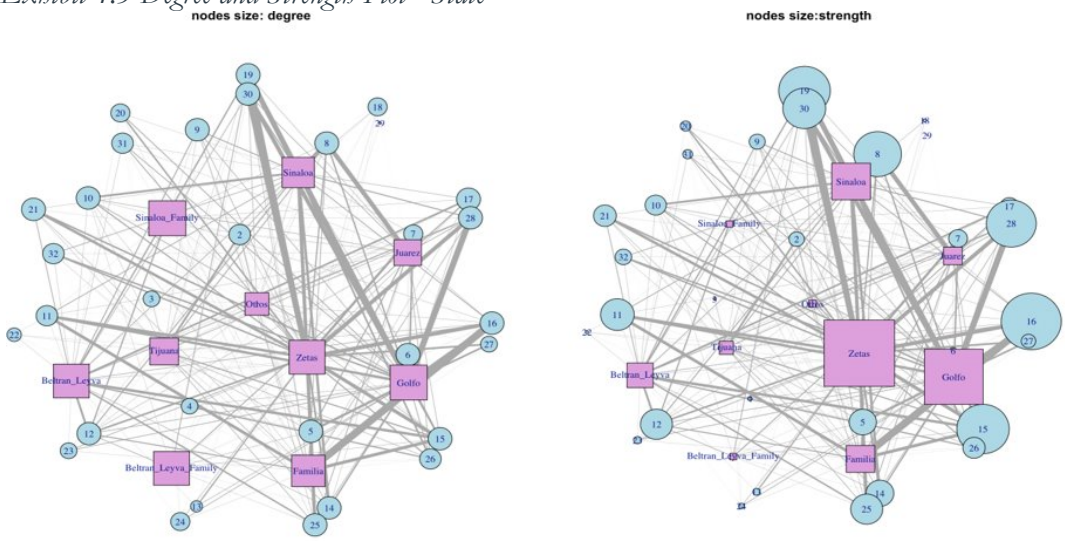
	groups	stren
9	Zetas	1587
4	Golfo	1320
6	Sinaloa	874
3	Familia	644
1	Beltran_Leyva	588
5	Juarez	421
8	Tijuana	316
10	Otros	171
2	Beltran_Leyva_Family	154
7	Sinaloa_Family	153

Table3

	states	stren
16	16	541
15	15	472
19	19	464
28	28	449
8	8	429
30	30	385
11	11	303
12	12	288
25	25	288
14	14	251
5	5	246
21	21	205
26	26	201
10	10	191
7	7	169

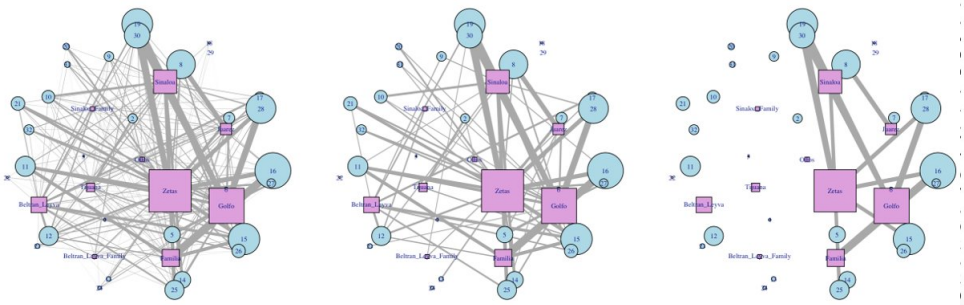
Table4

Exhibit 4.3 Degree and Strength Plot - State



We next performed weight analysis. Exhibit 4.3 shows all three plots for different edge weight: original edge weight, edge weight above mean 22, and edge weight above 100. We can see that the route from State 16 to Familiar has the highest value of 220 weight, meaning that there were 220 drug activities that happened from in State 16, which was Familiar's favorite route.

Exhibit 4.4 Weight



from	to	weight
16	Familia	220
28	Golfo	177
19	Zetas	155
19	Golfo	152
30	Zetas	151
8	Juarez	133
30	Golfo	121
16	Zetas	120
8	Sinaloa	119
28	Zetas	119
15	Zetas	104
16	Golfo	102
25	Sinaloa	100
15	Familia	96
11	Familia	92
11	Zetas	83
21	Zetas	83
15	Beltran_Leyva	82
11	Golfo	79

City

There are a total number of 2467 nodes and 2436 edges. The density is 0.0008. Reciprocity is 1 and there are 1693 components, meaning there are 1693 (1+1692) subgroups in this network. Noticed that nodes, edges, and subgroups have a much larger value comparing to State. When plotting City, it looks extremely messy, as shown in exhibit 4.3. We have tried to reduce our nodes to 775 as shown in exhibit 4.4. It worked, but still no significant insights discovered.

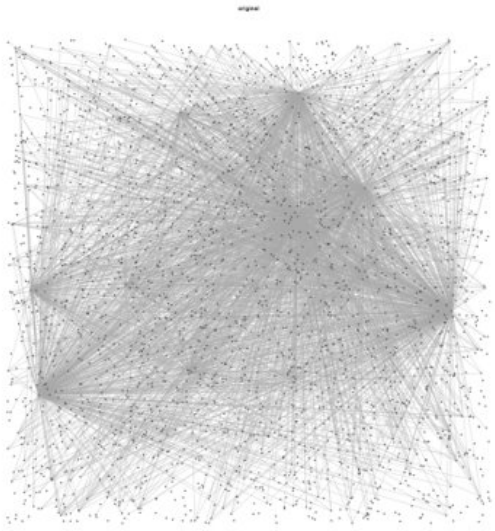


Exhibit 4.3 Original network of City

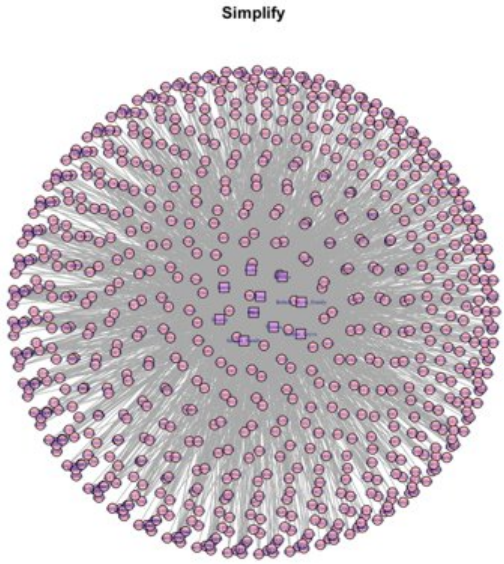


Exhibit 4.4 Simplified original network of City

Same as our State variable, we then added degree and strength analysis to City as well. Like our plots for our original network, not much information we can get. However, according to the matrix shown in exhibit 4.5.1, Zetas ranks first in degree with a number 585, meaning that Zetas has the highest spread, in terms of doing drug activities across 585 cities in Mexico, following by Golfo. Zetas also ranks first in the Strength analysis, with a number of 1587, meaning that Zetas has done drug activities 1587 times. Therefore, we can be able to conclude that Cartel Zetas had the widest, most spread out trafficking connection, and was the most active cartel.

Exhibit 4.5 Degree and Strength Plot – City

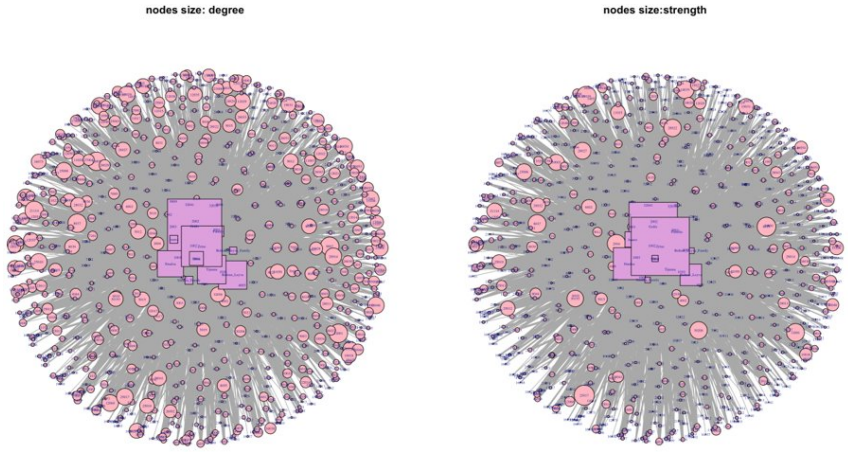


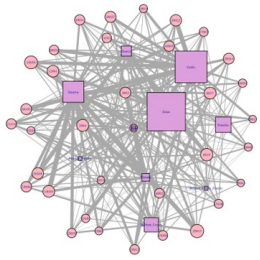
Exhibit 4.5.1 Matrix

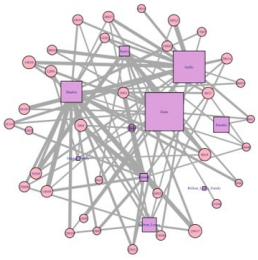
```
> sort(degree(dt2_S), decreasing = T)
             Zetas              Golfo           Familia           Sinaloa    Beltran_Leyva            Juarez           Tijuana
               585                438               308               307               282               151               115
Beltran_Leyva_Family             Otros    Sinaloa_Family              8037              9007             10005             12001
                86                 86                78                10                10                10                10
             21114              12029             12068             15027             15037             15051             15056
                10                  9                 9                 9                 9                 9                 9
             15106              16046             16054             17007             19039             25006             25010
                 9                  9                 9                 9                 9                 9                 9
```

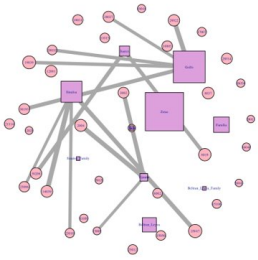
```
> sort(strength(dt2_S), decreasing = T)
             Zetas              Golfo           Sinaloa           Familia    Beltran_Leyva            Juarez           Tijuana
              1587               1320               874               644               588               421               316
             Otros Beltran_Leyva_Family    Sinaloa_Family             25017             19039              8019              2004
               171                154               153                44                43                41                40
              8037              12001             14039             28022             30208             28027              2002
                39                 39                39                38                38                37                36
             28014              30193             10005             25006             19031             21114              6002
                36                 36                35                35                34                34                32
```

Exhibit 4.6 shows all three plots for different edge weight: original edge weight, edge weight above mean 4.1, and edge weight above 10. From the plot and matrix, we see that the highest value of weight is 17 from City code 14039 to Sinaloa, meaning city code number 14039 is Sinaloa's favorite route. This route had happened 17 times. According to our data, different cartels use different routes. Tijuana, for example, heavily favored city code 02004.

Exhibit 4.6 Weight



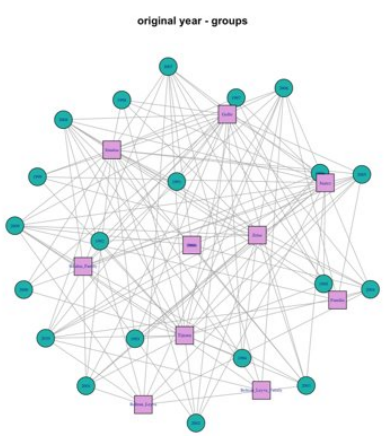


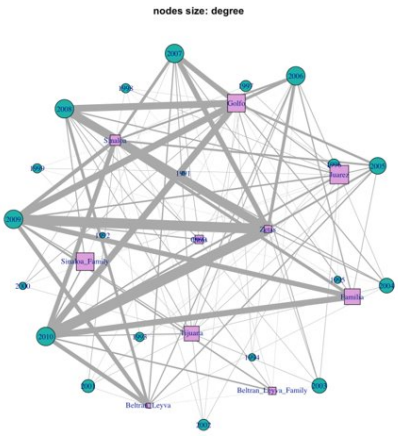


```
> busy_line[order(busy_line$weight,decreasing=TRUE),
    from             to weight
31 14039        Sinaloa     17
8   2004        Tijuana     16
76 28022          Golfo     16
3   2002        Tijuana     15
67 25017        Sinaloa     15
84 30193          Golfo     14
87 30208         Juarez     14
14  8019         Juarez     13
50 19039          Golfo     12
82 28032          Golfo     12
1   2001        Tijuana     11
60 25010        Sinaloa     11
79 28027          Golfo     11
34 15005        Tijuana     10
57 25006        Sinaloa     10
13  6002        Sinaloa      9
15  8019        Sinaloa      9
32 14039        Tijuana      9
46 19018          Golfo      9
53 21114          Golfo      9
62 25012        Sinaloa      9
```

Year

There are a total number of 31 nodes and 132 edges for year. Density is 0.28 and reciprocity equals 1. There are 2 components, meaning that there are 2 subgroups in this network. Plots in exhibit 4.7 shows that all 10 cartels have long drug activity history. We then added degree and strength to our network, the results are shown in exhibit 4.7.1 – 4.7.3.





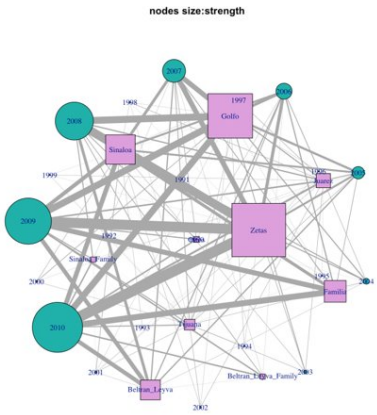


Exhibit 4.7 original *Exhibit 4.7.1 Degree* *Exhibit4.7.2 Strength*

According to our visualizations, we see that Cartel Golfo ranks first in degree with 20, meaning that Golfo has done drug activities across all 20 years in our data. Another example would be, Sinaloa ranks second, with a 19-year record of drug activities. On the other hand, Zetas ranks first when strength is added, meaning that Zetas has done the most drug activities, 1587 times throughout all 20 years.

4.7.3 Matrix

```
> sort(degree(dt3), decreasing = T)
```

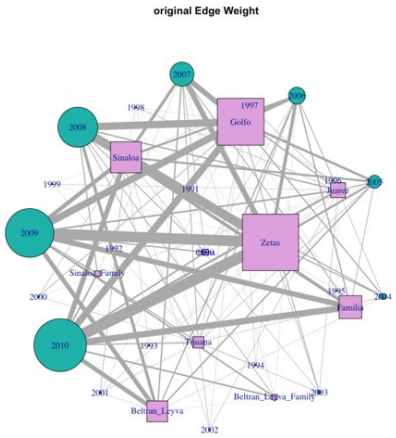
Golfo	Sinaloa	Tijuana	Juarez	Zetas	Sinaloa_Family	2006
20	19	19	17	16	11	10
2007	2008	2009	2010	2005	Beltran_Leyva	2003
10	10	10	10	9	9	8
2004	Familia	Otros	2001	1996	1997	2002
8	8	8	7	6	6	6
1993	1998	1999	Beltran_Leyva_Family	1994	1995	2000
5	5	5	5	4	4	4
1992	1991	1990				
3	2	0				

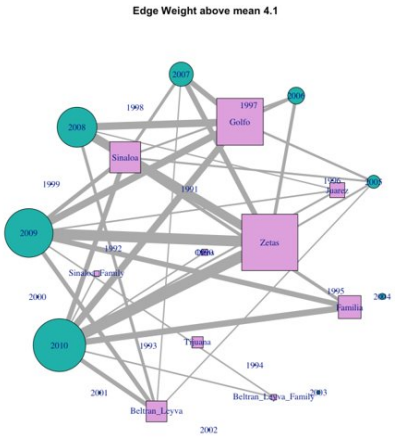
```
> sort(strength(dt3), decreasing = T)
```

Zetas	2010	2009	Golfo	2008	Sinaloa	2007
1587	1486	1369	1320	1126	874	680
Familia	Beltran_Leyva	2006	Juarez	2005	Tijuana	2004
644	588	476	421	367	316	171
Otros	Beltran_Leyva_Family	Sinaloa_Family	2003	2001	2000	2002
171	154	153	108	72	60	56
1997	1998	1996	1994	1999	1993	1995
51	49	36	31	31	24	23
1991	1992	1990				
6	6	0				

Exhibit*4.7.4* shows result after we added weight. According to our result, we see that the highest number of drug activity that had happened in one year is 405, in year 2010 by Zetas. We can also interpret from the result, for example, that Zetas was the most active cartel from year 2008 to 2010.

Exhibit 4.7.4 Weight






```
> busy_line[order(busy_line$weight,decreasing=TRUE),]
    from                to weight
131 2010             Zetas    405
121 2009             Zetas    387
111 2008             Zetas    321
107 2008             Golfo    265
117 2009             Golfo    254
127 2010             Golfo    244
125 2010           Familia    227
115 2009           Familia    192
97  2007             Golfo    190
101 2007             Zetas    190
128 2010           Sinaloa    176
113 2009      Beltran_Leyva    157
123 2010      Beltran_Leyva    157
118 2009           Sinaloa    147
108 2008           Sinaloa    142
91  2006             Zetas    119
105 2008           Familia    118
87  2006             Golfo    112
98  2007           Sinaloa    109
103 2008      Beltran_Leyva    108
77  2005             Golfo     86
```

Centrality analysis

We have also performed a centrality analysis, as shown in exhibit 5. Like degree and strength, betweenness and closeness are alternative measurements that provide insights. Betweenness centrality is a way of detecting the amount of influence a node has and Closeness is also a way of detecting nodes, that are able to spread information in a efficiency way. Since we are using affiliation network, which focusing on the relationship between Cartels and State/City, degree analysis provides greater information then centrality analysis.

Exhibit 5 Centralization

```
> dt1.promsort2
                        degree closeness betweenness
Sinaloa              31.000000 0.7884615  68.0896746
Zetas                31.000000 0.7884615 102.1429293
Beltran_Leyva        30.000000 0.7592593  62.1429293
Familia              30.000000 0.7592593  63.3271473
Golfo                30.000000 0.7592593  63.0650332
Juarez               28.000000 0.7068966  50.6448692
Tijuana              27.000000 0.6833333  46.6454107
Otros                24.000000 0.6212121  35.3479497
Sinaloa_Family       23.000000 0.6029412  29.3174195
Beltran_Leyva_Family 20.000000 0.5540541  22.2766372
5                    10.000000 0.5694444   3.9507256
6                    10.000000 0.5694444   3.9507256
8                    10.000000 0.5694444   3.9507256
9                    10.000000 0.5694444   3.9507256
10                   10.000000 0.5694444   3.9507256
12                   10.000000 0.5694444   3.9507256
14                   10.000000 0.5694444   3.9507256
15                   10.000000 0.5694444   3.9507256
16                   10.000000 0.5694444   3.9507256
17                   10.000000 0.5694444   3.9507256
Centralization        0.437863 0.4222470   0.1087151
```

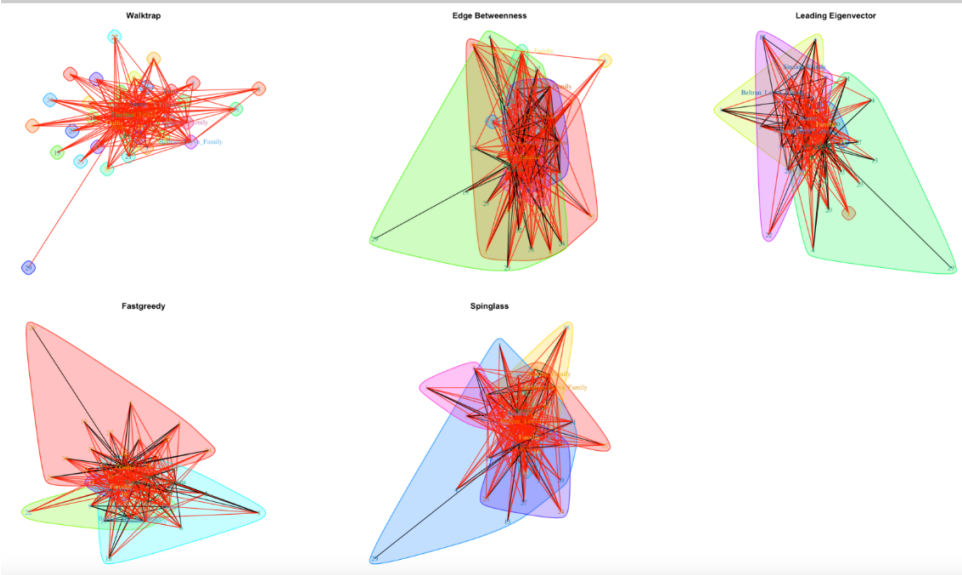
```
> dt2.promsort2
                        degree closeness  betweenness
Zetas                585.0000000 0.6771654 1.481462e+05
Golfo                438.0000000 0.5386221 6.337702e+04
Familia              308.0000000 0.4560990 5.207908e+04
Sinaloa              307.0000000 0.4555621 4.084063e+04
Beltran_Leyva        282.0000000 0.4425386 2.925011e+04
Juarez               151.0000000 0.3848831 8.003403e+03
Tijuana              115.0000000 0.3715795 3.825097e+03
Beltran_Leyva_Family  86.0000000 0.3615133 3.008361e+03
Otros                 86.0000000 0.3615133 3.098737e+03
Sinaloa_Family        78.0000000 0.3588317 2.218378e+03
8037                  10.0000000 0.5032510 4.422591e+02
9007                  10.0000000 0.5032510 4.422591e+02
10005                 10.0000000 0.5032510 4.422591e+02
12001                 10.0000000 0.5032510 4.422591e+02
21114                 10.0000000 0.5032510 4.422591e+02
12029                  9.0000000 0.5019455 4.069264e+02
12068                  9.0000000 0.5012953 3.971700e+02
15027                  9.0000000 0.4771887 2.680453e+02
15037                  9.0000000 0.5019455 4.069264e+02
15051                  9.0000000 0.5025974 3.881879e+02
Centralization         0.7476919 0.4998719 4.940674e-01
```

```
> dt3.promsort2
                  degree closeness betweenness
Golfo         20.0000000 0.4347826  63.4775079
Sinaloa       19.0000000 0.4225352  53.4614880
Tijuana       19.0000000 0.4225352  49.4286346
Juarez        17.0000000 0.4000000  34.4201063
Zetas         16.0000000 0.3896104  32.1606282
Sinaloa_Family 11.0000000 0.3448276  10.5352208
2006          10.0000000 0.3797468  14.2099372
2007          10.0000000 0.3797468  14.2099372
2008          10.0000000 0.3797468  14.2099372
2009          10.0000000 0.3797468  14.2099372
2010          10.0000000 0.3797468  14.2099372
2005           9.0000000 0.3703704   9.4099372
Beltran_Leyva  9.0000000 0.3296703   5.3685075
2003           8.0000000 0.3614458   6.7258225
2004           8.0000000 0.3614458   7.4613939
Familia        8.0000000 0.3225806   4.3796110
Otros          8.0000000 0.3225806   3.7682957
2001           7.0000000 0.3529412   4.3285199
1996           6.0000000 0.3448276   2.4214633
1997           6.0000000 0.3448276   3.5374361
Centralization 0.3827957 0.1855321   0.1223604
```

community detection

Our community detection results and plots are shown in Exhibit 6. Our analysis indicates that our network is either widely spreading or overlapping. There is no obvious community detection because drug activities cover nearly every single State/City.

Exhibit 6 Community Detection



- modularity(cw): 0
- modularity(ceb): 0.007368287
- modularity(cle): 0.2294019
- modularity(cfg):0.23126
- modularity(cs): 0.006596966

Modeling

ERGM

Because we do not have enough attributes, we decided to add two attributes to our model – "Size" and Color". Process shows in exhibit 7.

Exhibit 7 Modeling

```
ergm_model <- ergm(dt1_n ~ edges + nodefactor('color') + nodecov('size') )
```

Notes: add color and size two attributes

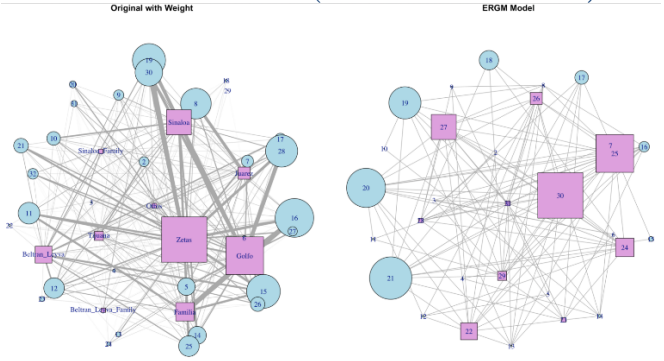
- Model fitness

```
PDSmod.fit <- gof(ergm_model, GOF = ~ distance + espartners + degree
+ triadcensus, burnin=1e+5, interval = 1e+5)
```

Monte Carlo MLE Results:

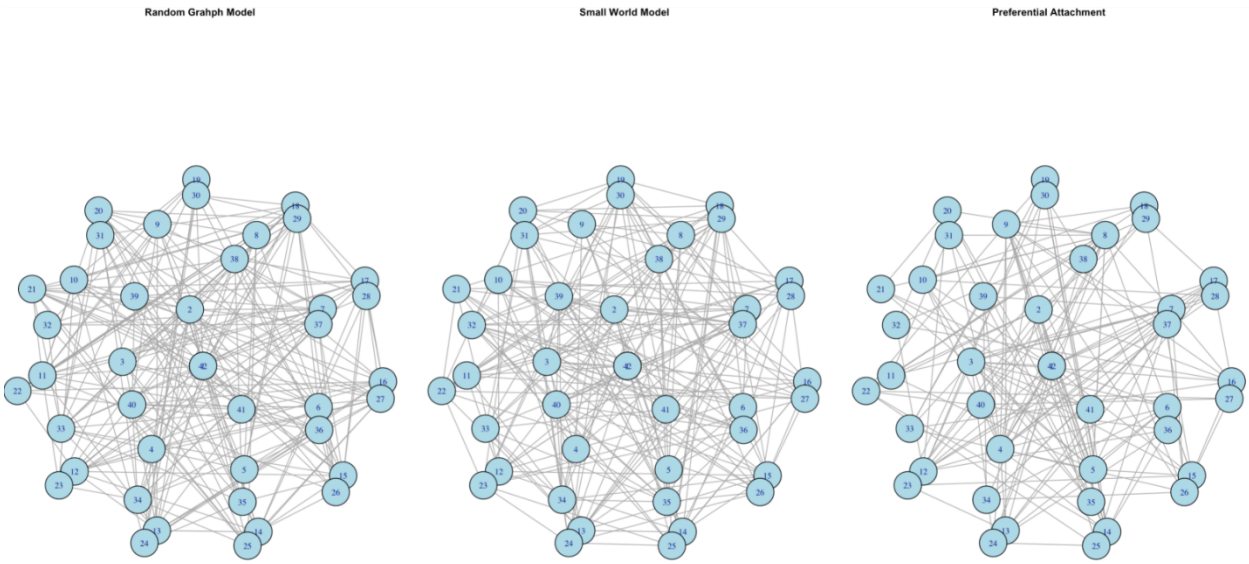
	Estimate	Std. Error	MCMC %	z value	Pr(>|z|)	
edges	-2.52261	0.23943	0	-10.536	<1e-04	***
nodefactor.color.plum	0.99943	0.18207	0	5.489	<1e-04	***
nodecov.size	0.04733	0.00866	0	5.466	<1e-04	***

AIC: 474.1 BIC: 486.5 (the smaller the better.)



Small world

```
gs0 <- watts.strogatz.game(dim=1, size=42, nei=5, p=0.25)
gs1 <- watts.strogatz.game(dim=1, size=42, nei=5, p=1)
gp <- barabasi.game(42, out.dist = c(0,0.01,0.01, 0.01,0.70), directed=FALSE, zero.appeal=1)
```



ERGM, Small world, and random graph can capture information such as the count of nodes size, nodes color and edge number. Due to the fact that we only have a few attributes. The above models only tell one part of the story, meaning that they provides limited information.

6. Other relevant methodological analyses

Tableau

Tableau is one of the alternatives that is better at visualizations. Tableau is good at creating various map such as Heat map, Flow maps, Spider maps, Point, distribution maps, and Filled map on each of our three variables. Tableau can also compare the drug traffic groups using different bar charts. What is more, is that Tableau can analysis the year trending through the line graph, which gives a simple, yet direct visualizations on the activities throughout 20 years. However, Tableau is not the ideal system for finding the relationships between drug traffic group and Location. As we have mentioned, Tableau is more towards visualization. R is preferable if interpretation is prioritized.

Excel

Both R and Excel are excellent data analytics tools, but they each have distinct functionality. Excel is a well-known software program that created by Microsoft. User can create spreadsheets, execute calculations, produce charts, and perform statistical analysis. Excel has been used in mostly every industry.

Most people nowadays do have a basic knowledge on how to use excel to complete their work. The benefit of excel is the initial learning curve is quite minimal, and most analysis can be done via point-and click on the top panel. However, the initial learning curve of R is steeper. It will take most at least a few weeks to familiarize themselves with the interface and master the various

functions. Next, we will use visualization as an example to describe the difference between R and Excel.

Using R or Excel as your visualization tool that is depended on your needed. You can ask yourself, "How detailed do my visualizations need to be in order to achieve my goal(s)?" In Excel, user can easily perform basic analyst of the dataset. such as count, highlight a group of cells, and make a simple chart for PowerPoint. Exhibit 6 shows the chart we created based on the year dataset. From this chart, we can easily and quickly obtain some basic information. However, if detailed information, algorisms, or complicated statics is needed, Excel is lacking in the perspective of of its functionality. R may be the best bet. R can produce incredibly attractive, detailed visuals and as well as detailed information, which is extremely helpful for understanding and interpreting.

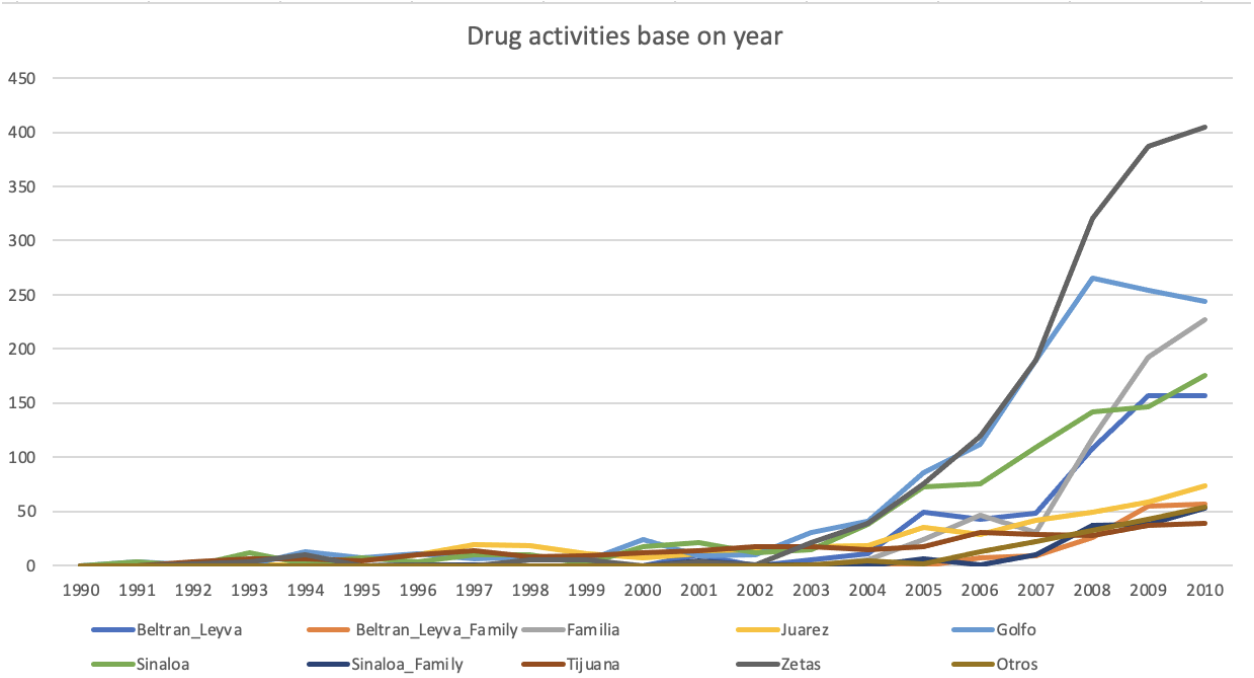


Exhibit 6 Example of Excel graph (Drug Activities base on Year)

7. Result & Discussion

In this study, we have successfully built a framework using R to see the association and patterns between Mexican municipalities and drug traffic organization. In addition, our model provides significant detailed information on 10 Cartels drug activities throughout 20 years of data. For better analysis, we first split our data into three small datasets - state, city, and year. We then performed network analysis using R on 5 number summary, degree, strength, and weight on each 3 variables. In addition, we have also done analysis with centralization, community detection, and ERGM modeling. From our result, we can be able to conclude that our findings are contributing to the understanding of the drug trafficking routes in Mexico.

There are two limitations. First, our data only captured data from 1990 – 2010, which is considerably outdated. Our result does not accurately reflect current situations. El Chapo, or "shorty" for example, was sentenced to life in prison in July 2019. However, his jailing led to an increase of

violence in the region of Sinaloa and Sinaloa Cartel remains hugely powerful. Therefore, even though our result does not accurately reflect current situations, it is still worth studying because valuable information that comes from the history.

Second limitation is reliability, as we have mentioned earlier. The data was extracted from Google news articles, which raise the concern of accuracy, or reliability. Based on the description of data and background literatures we found, data prior to 2004 is still reported, but less reliable than data after 2004. There are countless activities hidden behind that would never show on the news. Therefore, reliability is always a concern.

Lastly, there are many other factors that influencing the Mexico drug route selection. For example, politics and corruptions. Almost a well-known secret that corruptions of law enforcements in Mexico is one of the key factors that supporting the growth of drug cartels. Study found that political affiliation shifts in the administration of different municipalities influenced drug violence and diverging drug smuggling to alternative routes. However, these aspects are out of the scope of our study, but still weighted heavily to this topic.

Reference

[1] Medel, M., et al., Mexico's drug networks: Modeling the smuggling routes towards the United States, Applied Geography (2014), https://www.sciencedirect.com/science/article/abs/pii/S0143622814002550

[2] Mexico cartels: Which are the biggest and most powerful? BBC (2019) https://www.bbc.com/news/world-latin-america-40480405

[3] Michele Coscia, Virdiana Rios, Knowing Where and How Criminal Organizations Operate Using Web Content, Harvard (2012) https://projects.iq.harvard.edu/ptr/files/cosciarios.pdf

[4] Melissa Dell, Trafficking Networks and the Mexican Drug War, Harvard (2012) https://economics.columbian.gwu.edu/sites/g/files/zaxdzs1936/f/downloads/Trafficking%20Networks%20and%20the%20Mexican%20Drug%20War.pdf