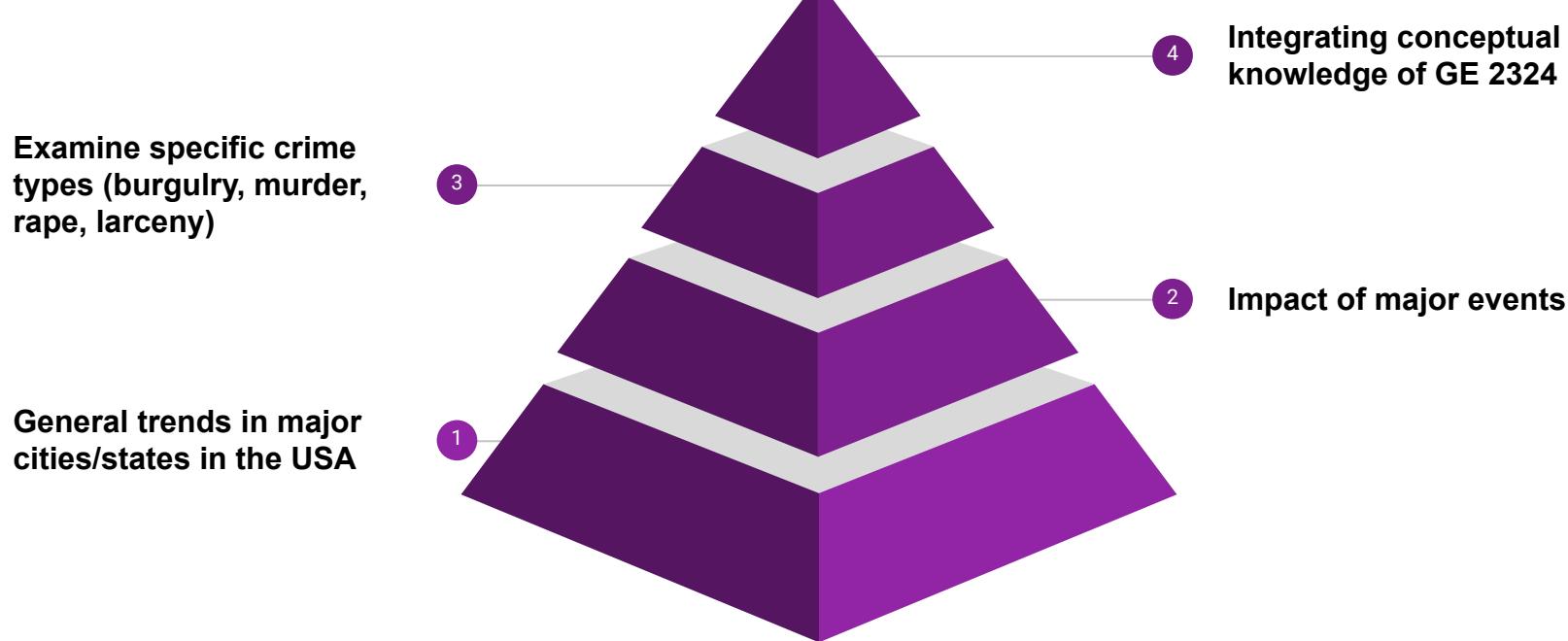


Analysis of Crime rates and Statistics in the United States

T03 Group A's Presentation

Project Introduction & Scope





Raw Data

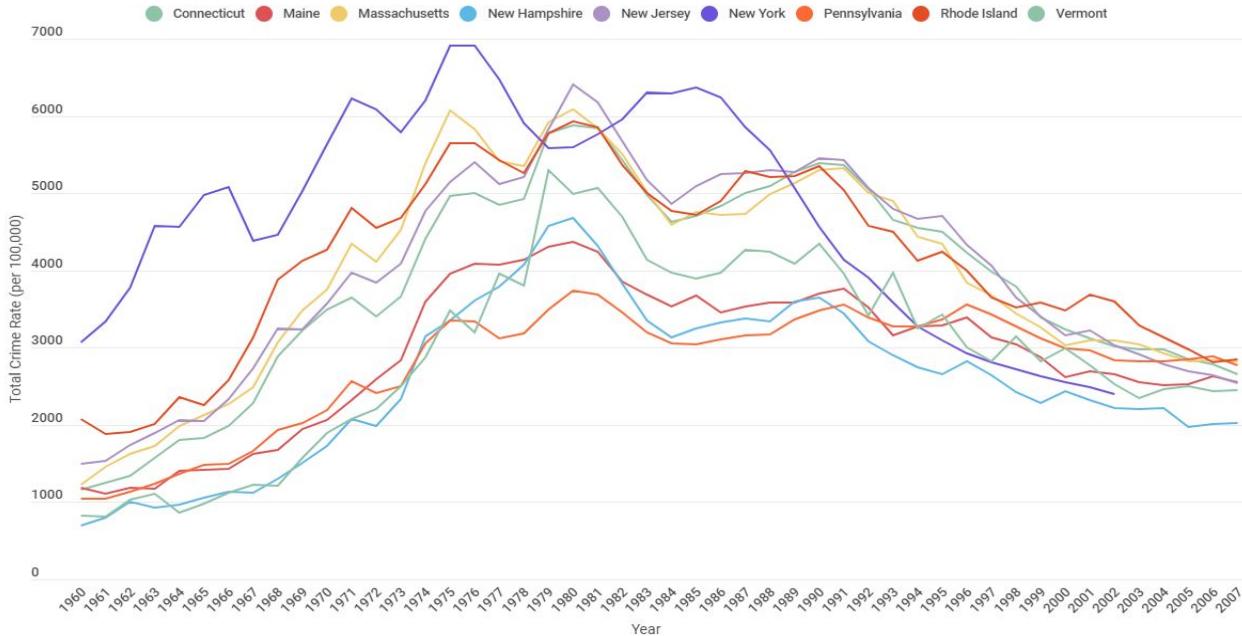
The US government's open data

Tally of all reported crimes in the
US for 1960-2019

Information about Police Precincts
in cities

Data logs of shooting incidents in
the NYPD

General trends in crimes in major cities

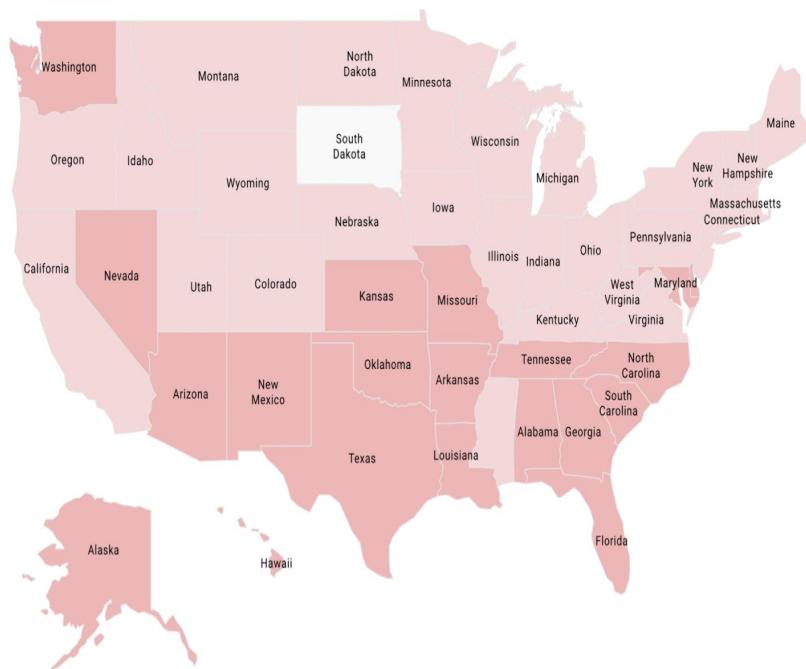


Key Features:

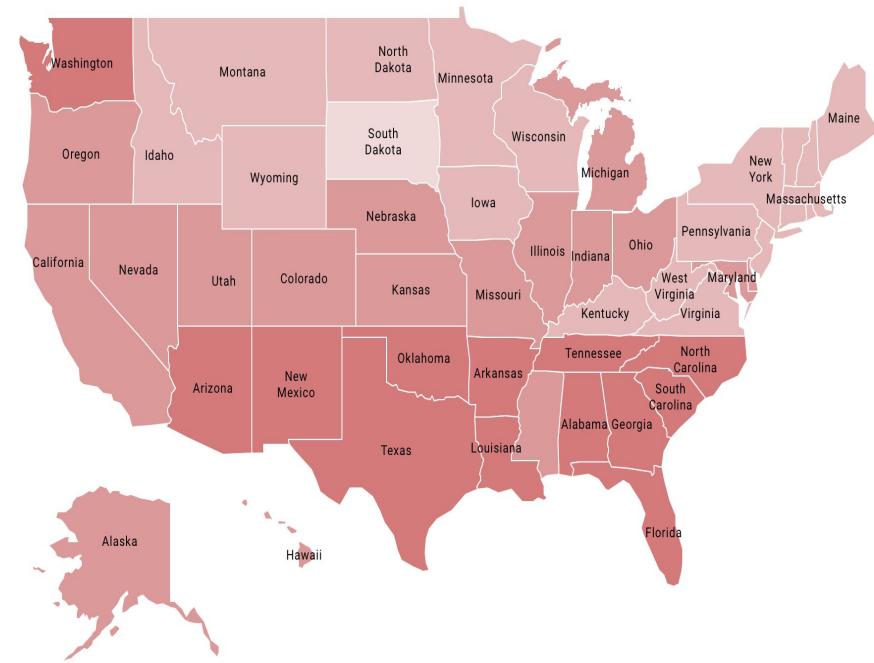
- New York having most no. of crime rates in year 1960
- Maine, Connecticut, New Jersey, and Massachusetts following the same trend
- Sudden decline from year 1990
- New York declined the most
- In year 2007 New Hampshire still having least crime rate

Impact of the 2008 stock market crash

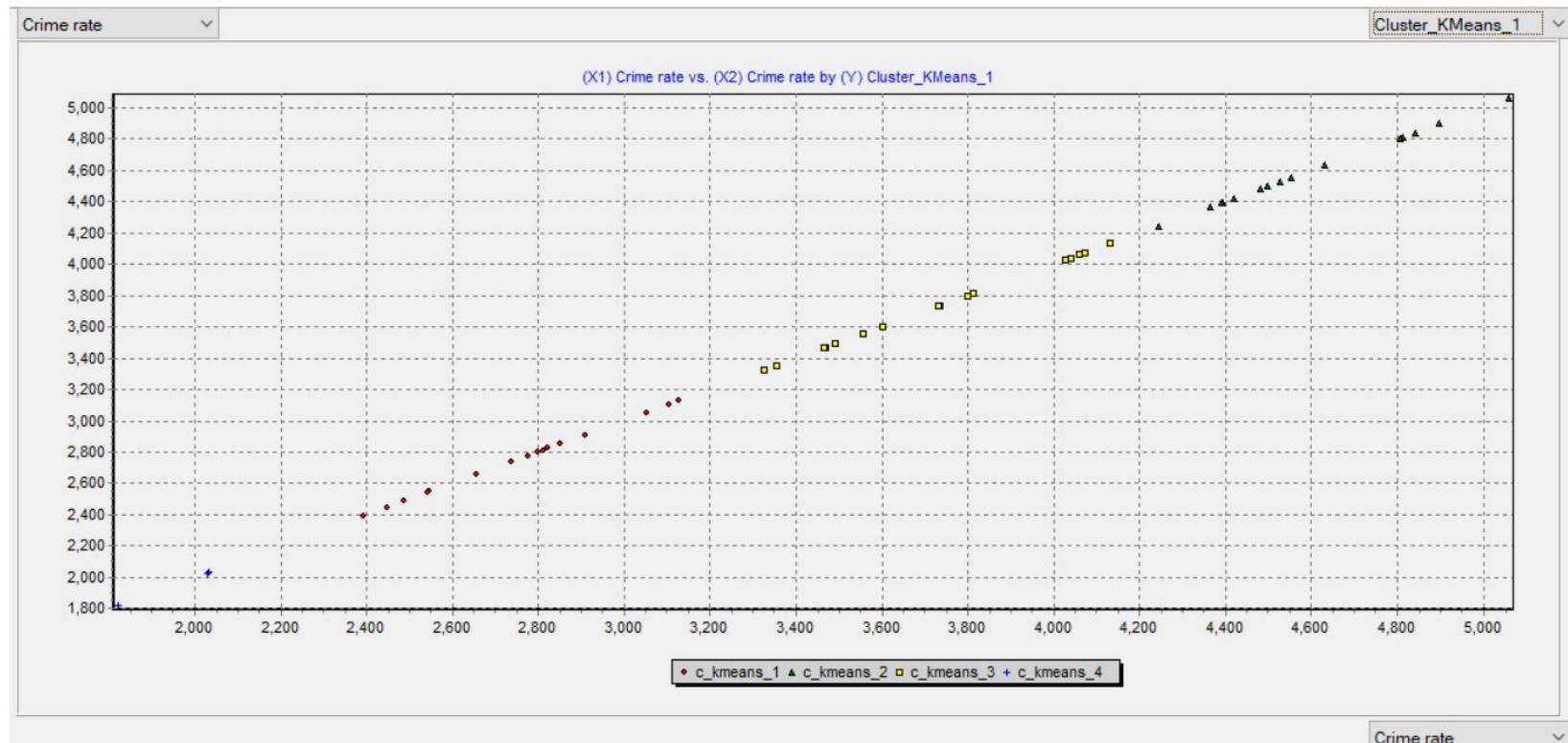
Year 2007

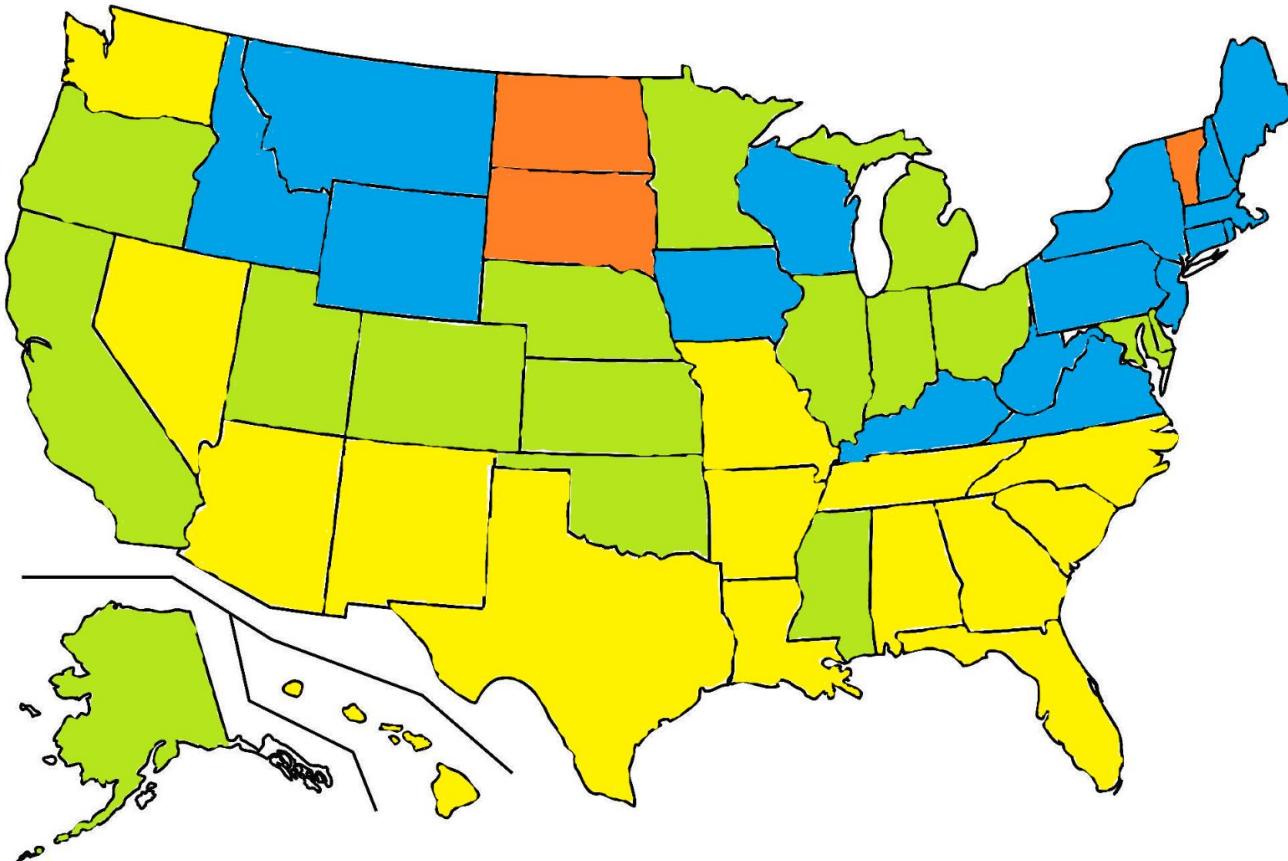


Year 2009



Data Clustering and Dendrogram - Crime Index Total (2007)





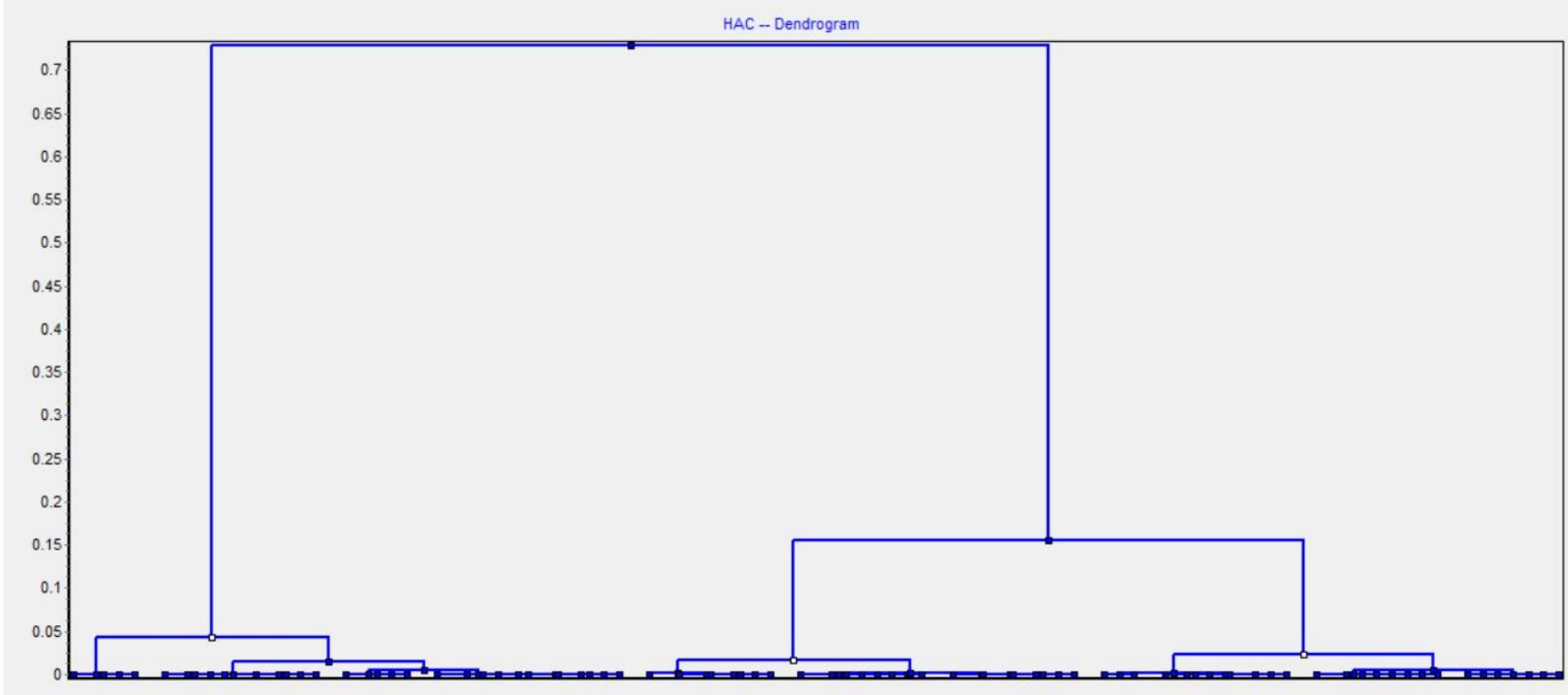
Yellow - Cluster 1
(Highest)

Green - Cluster 2

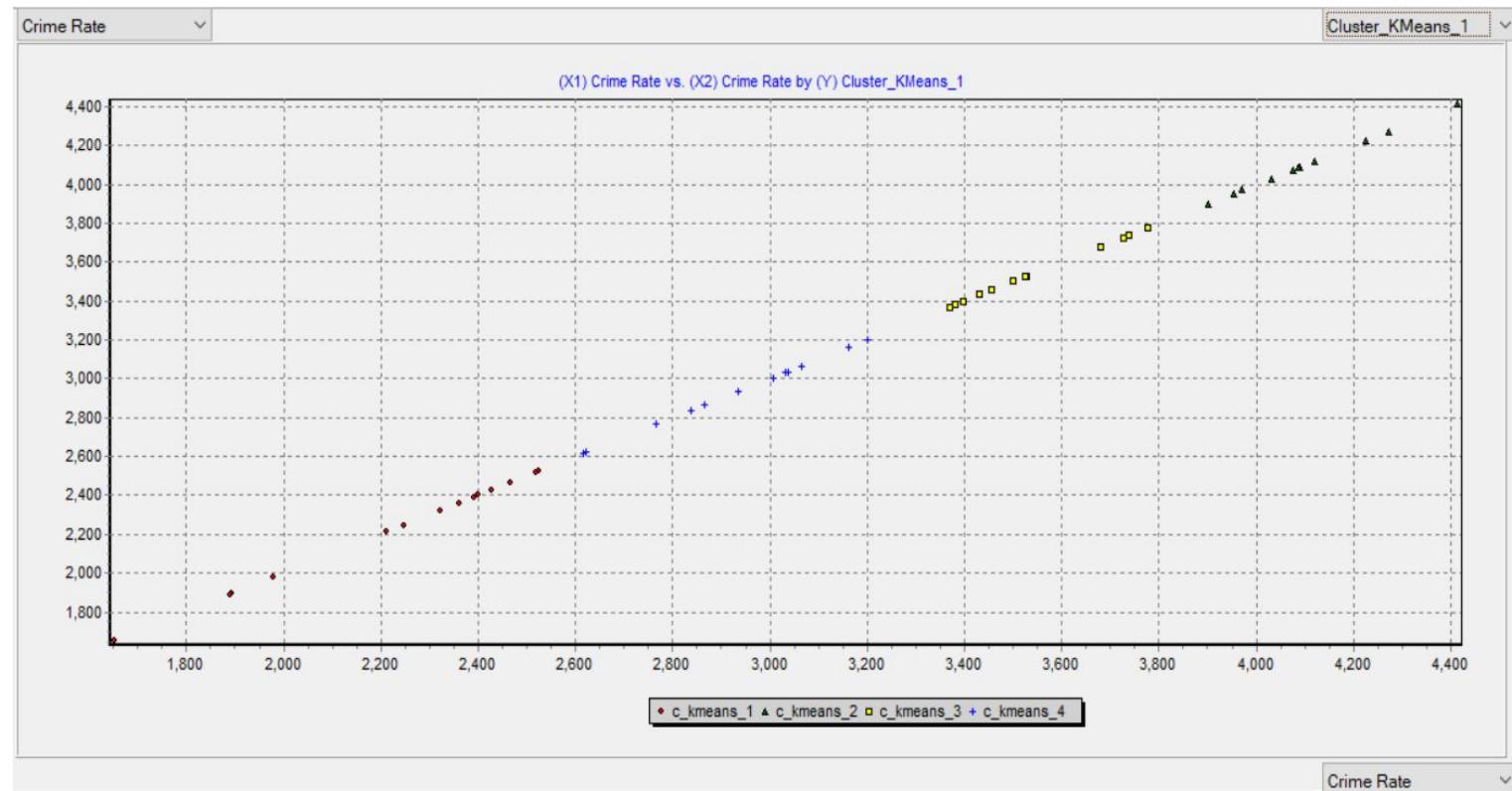
Blue - Cluster 3

Orange - Cluster 4
(Lowest)

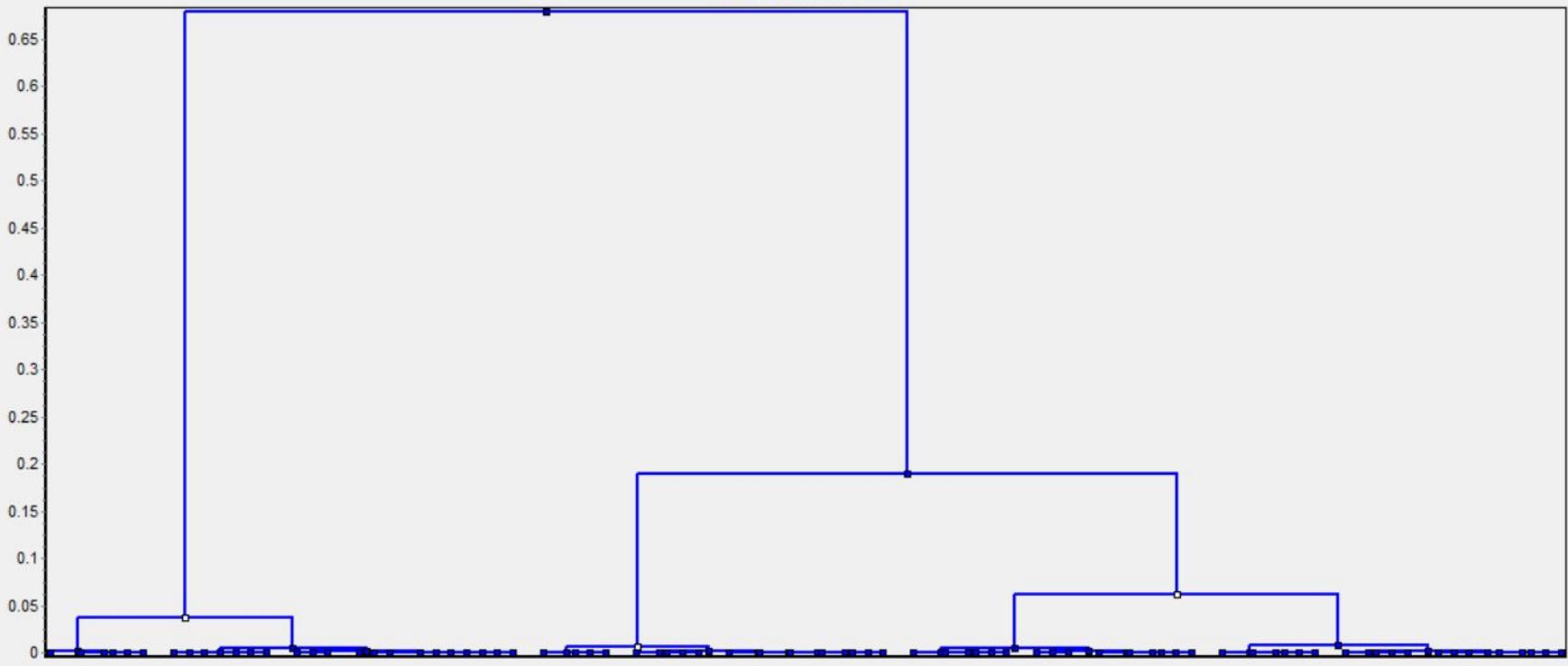
HAC -- Dendrogram



Data Clustering and Dendrogram - Property Crime (2007)



HAC -- Dendrogram



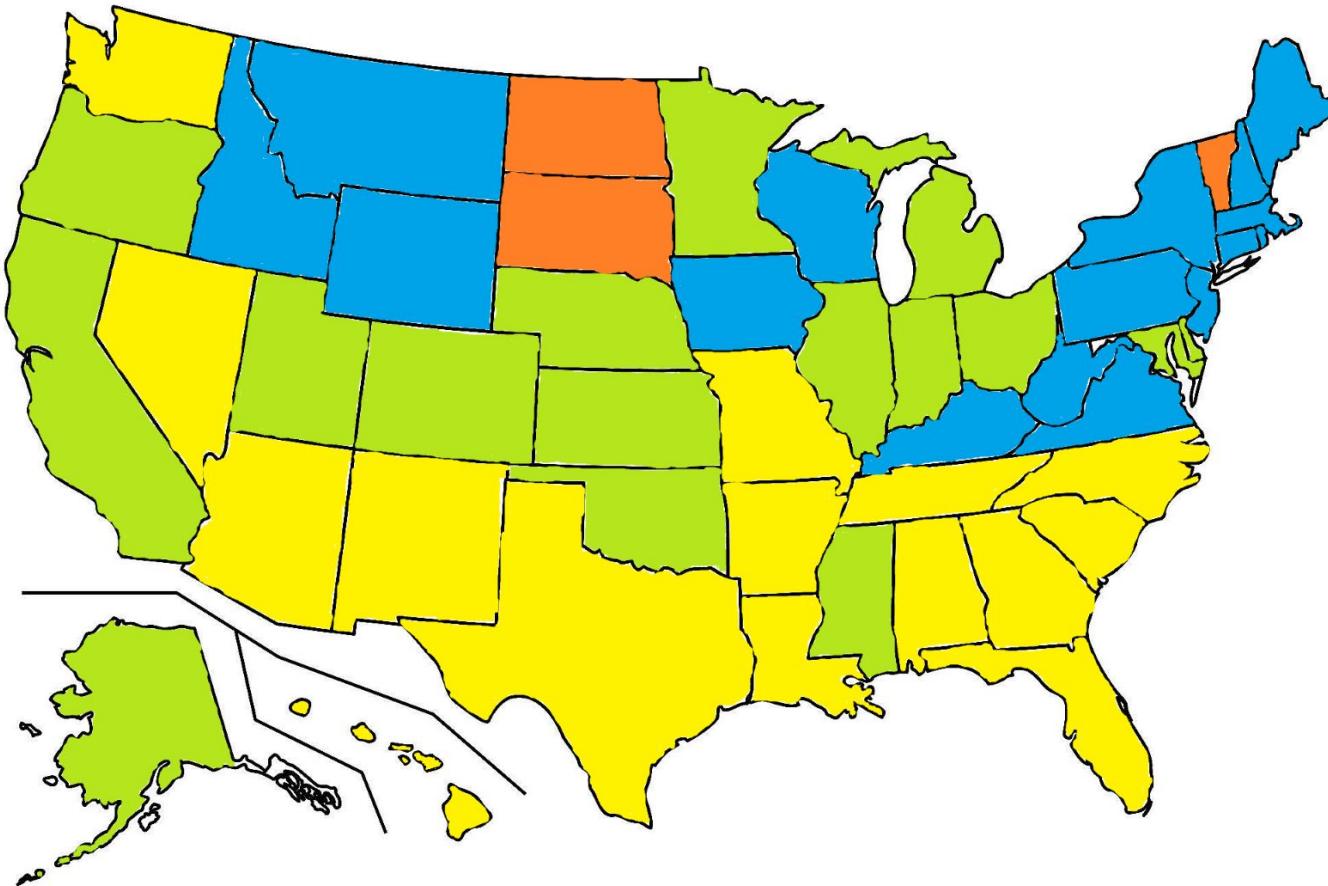
Analysis Based on the Crime Rates

Low Crime Rates

- New Hampshire
- Vermont
- North Dakota

High Crime Rates

- Arizona
- South Carolina



Yellow - Cluster 1
(Highest)

Green - Cluster 2

Blue - Cluster 3

Orange -Cluster 4
(Lowest)

Association Between Different Crimes in US

- Association is found using Apriori Algorithm
- Requirement for Apriori Algorithm:
 - A table filled with true/false, explaining the features of records

What we had

Year	State	Murder and nonnegligent manslaughter rate	Forcible rape rate	Robbery rate	Aggravated assault rate	Property crime rate	Burglary rate	Larceny-theft rate	Motor vehicle theft rate
1999	United States	5.7	32.8	150.1	334.3	3743.6	770.4	2550.7	422.5
2000	United States	5.5	32.0	145.0	324.0	3618.3	728.8	2477.3	412.2
2001	United States	5.6	31.8	148.5	318.5	3656.1	740.8	2484.6	430.6
2002	United States	5.6	33.1	146.1	309.5	3630.6	747.0	2450.7	432.9
2003	United States	5.7	32.3	142.5	295.4	3591.2	741.0	2416.5	433.7
2004	United States	5.5	32.4	136.7	288.6	3514.1	730.3	2362.3	421.5
2005	United States	5.6	31.8	140.8	290.8	3431.5	726.9	2287.8	416.8
2006	United States	5.7	31.0	149.4	287.5	3334.5	729.4	2206.8	398.4
2007	United States	5.6	30.0	147.6	283.8	3263.5	722.5	2177.8	383.3
AVERAGE		7.5	28.9	175.8	270.4	3946.3	1050.3	2452.9	443.1

What we Made

Year	Country	Murder and nonnegligent manslaughter rate	Forcible rape rate	Robbery rate	Aggravated assault rate	Property crime rate	Burglary rate	Larceny-theft rate	Motor vehicle theft rate
1999	United States	false	true	false	true	false	false	true	false
2000	United States	false	true	false	true	false	false	true	false
2001	United States	false	true	false	true	false	false	true	false
2002	United States	false	true	false	true	false	false	false	false
2003	United States	false	true	false	true	false	false	false	false
2004	United States	false	true	false	true	false	false	false	false
2005	United States	false	true	false	true	false	false	false	false
2006	United States	false	true	false	true	false	false	false	false
2007	United States	false	true	false	true	false	false	false	false

Association Between Different Crimes in US

Association Between Different Crimes in US

Association Rules

1. Aggravated assault rate=TRUE 29 ==> Forcible rape rate=TRUE 29 <conf:(1)>
2. Property crime rate=TRUE 25 ==> Larceny-theft rate=TRUE 25 <conf:(1)>
3. Burglary rate=TRUE 24 ==> Murder and nonnegligent manslaughter rate=TRUE 24 <conf:(1)>
4. Robbery rate=TRUE Larceny-theft rate=TRUE 24 ==> Property crime rate=TRUE 24 <conf:(1)>
5. Robbery rate=TRUE Property crime rate=TRUE 24 ==> Larceny-theft rate=TRUE 24 <conf:(1)>
6. Murder and nonnegligent manslaughter rate=TRUE 26 ==> Robbery rate=TRUE 25 <conf:(0.96)>
7. Property crime rate=TRUE 25 ==> Robbery rate=TRUE 24 <conf:(0.96)>
8. Property crime rate=TRUE Larceny-theft rate=TRUE 25 ==> Robbery rate=TRUE 24 <conf:(0.96)>
9. Property crime rate=TRUE 25 ==> Robbery rate=TRUE Larceny-theft rate=TRUE 24 <conf:(0.96)>
10. Forcible rape rate=TRUE 31 ==> Aggravated assault rate=TRUE 29 <conf:(0.94)>

Most Interesting Ones

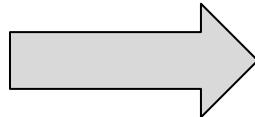
3. Burglary rate=TRUE 24 ==> Murder and nonnegligent manslaughter rate=TRUE 24 <conf:(1)>
4. Robbery rate=TRUE Larceny-theft rate=TRUE 24 ==> Property crime rate=TRUE 24 <conf:(1)>
5. Robbery rate=TRUE Property crime rate=TRUE 24 ==> Larceny-theft rate=TRUE 24 <conf:(1)>

Association Between Different Crimes in US

Most Interesting Ones



Burglary Rate



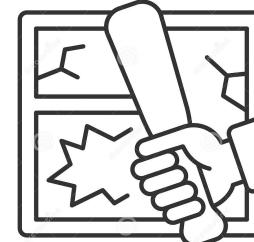
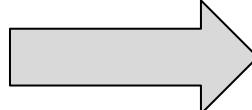
Murder and nonnegligent manslaughter Rate



Robbery Rate

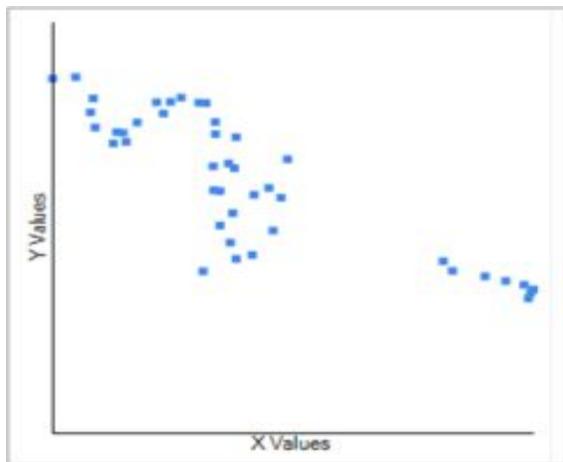


Larceny Rate



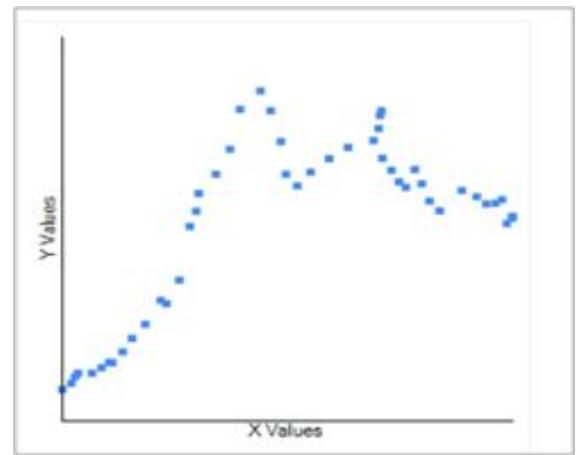
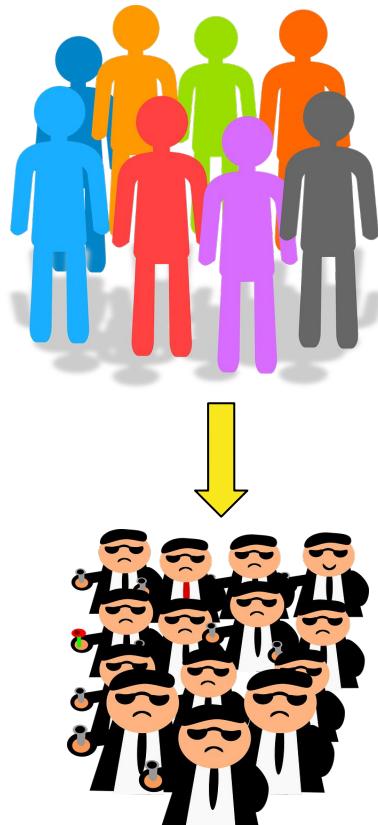
Property Crime Rate

Correlation between increasing population and reported crimes in New York and New Jersey



New York

Strong Negative Correlation
-0.839



New Jersey

Moderately Positive Correlation
0.676

Assumptions on the result

New York

- Substantial decrease in reported crime with an increase in population over 42 years.
- More number of violent crimes and assaults than thefts and burglaries in the period.
- Does not agree with the general trend of population and reported crime.



New Jersey

- Increase in crime with an increase in population over 47 years.
- More number of thefts and burglaries than assaults and violent crimes in the period.
- Agrees with the general trends of population and reported crime.

Reasons for the irregularities

New York

- Increase in the number of street gangs during the 80s.
- The crack or the cocaine epidemic during the 70s.
- Insufficient state police budget during the 90s.
- Massive increase in security of the state post 9/11.

New Jersey

- Lower number of aggressive assaults than New York but high in theft and burglary.
- Not much change in the number of deployed police officials on the street post 9/11.
- Not a financial hub like New York.



Pearson correlation plot on various var (1977 data of 50 States)

	Population	Illiteracy	Life Exp	Murder	HS Grad	Frost
Population	1.00000000	0.1076224	-0.06805195	0.3436428	-0.09848975	-0.3321525
Illiteracy	0.10762237	1.0000000	-0.58847793	0.7029752	-0.65718861	-0.6719470
Life Exp	-0.06805195	-0.5884779	1.00000000	-0.7808458	0.58221620	0.2620680
Murder	0.34364275	0.7029752	-0.78084575	1.0000000	-0.48797102	-0.5388834
HS Grad	-0.09848975	-0.65718861	0.58221620	-0.4879710	1.00000000	0.3667797
Frost	-0.33215245	-0.6719470	0.26206801	-0.5388834	0.36677970	1.0000000

2018 data of 50 States

	States	total_area	unemployment	population	murder	annual_rainfall_inches	literacy_rate
1	Alabama	135767	3.9	5024279	58.3	0.852	
2	Alaska	1723337	6.6	733391	180000	22.5	0.908
3	Arizona	295234	4.8	7151502	33141	13.6	0.869
4	Arkansas	137732	3.7	3011524	17643	50.6	0.863
5	California	423967	4.2	39538223	174331	22.2	0.769
6	Colorado	269601	3.3	5773714	21938	15.9	0.901
7	Connecticut	14357	4.1	3605944	6546	50.3	0.914
8	Delaware	6446	3.8	989948	4115	45.7	0.893
9	Florida	170312	3.6	21538187	81270	54.5	0.803
10	Georgia	153910	3.9	10711908	36170	50.7	0.833
11	Hawaii	28313	2.4	1455271	4042	63.7	0.841
12	Idaho	216443	2.8	1839106	4000	18.9	0.895
13	Illinois	149995	4.3	12812508	51561	39.2	0.871
14	Indiana	94326	3.4	6785528	24966	41.7	0.920
15	Iowa	145746	2.5	3190369	8410	34.0	0.925
16	Kansas	213100	3.4	2937880	11968	28.9	0.922
17	Kentucky	104656	4.3	4505836	9701	48.9	0.878
18	Louisiana	135659	4.9	4657757	25537	60.1	0.840
19	Maine	91633	3.4	1362359	1548	42.2	0.926
20	Maryland	32131	3.9	6177224	27456	44.5	0.888
21	Massachusetts	27336	3.3	7029917	22578	47.7	0.901
22	Michigan	250487	4.1	10077331	43686	32.8	0.917
23	Minnesota	225163	2.9	5706494	13332	27.3	0.940
24	Mississippi	125438	4.8	2961279	8272	59.0	0.840
25	Missouri	180540	3.2	6154913	30380	42.2	0.925
26	Montana	380831	3.7	1084225	4328	15.3	0.912
27	Nebraska	200330	2.8	1961504	5821	23.6	0.927

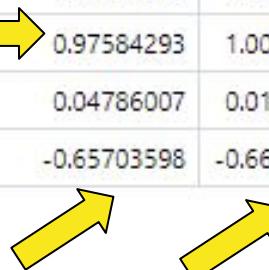
Showing 1 to 27 of 50 entries, 7 total columns

	States	total_area	unemployment	population	murder	annual_rainfall_inches	literacy_rate
26	Montana	380831	3.7	1084225	4328	15.3	0.912
27	Nebraska	200330	2.8	1961504	5821	23.6	0.927
28	Nevada	286380	4.6	3104614	15210	9.5	0.839
29	New Hampshire	24214	2.5	1377529	2074	43.4	0.942
30	New Jersey	22591	4.1	9288994	18375	47.1	0.831
31	New Mexico	314917	4.9	2117522	17450	14.6	0.835
32	New York	141297	4.1	20201249	69764	41.8	0.779
33	North Carolina	139391	3.9	10439388	38995	50.3	0.864
34	North Dakota	183108	2.6	779094	2169	17.8	0.937
35	Ohio	116098	4.6	11799448	34269	39.1	0.809
36	Oklahoma	181037	3.4	3959353	17086	36.5	0.877
37	Oregon	254799	4.2	4237256	11995	27.4	0.898
38	Pennsylvania	119280	4.3	13002700	39228	42.9	0.874
39	Rhode Island	4001	4.1	1097379	2342	47.9	0.915
40	South Carolina	82933	3.4	5118425	26323	49.8	0.853
41	South Dakota	199729	3.0	886667	3530	20.1	0.930
42	Tennessee	109153	3.5	6910840	40647	54.2	0.868
43	Texas	695662	3.9	29145505	121474	28.9	0.810
44	Utah	219882	3.1	3271616	7553	12.2	0.906
45	Vermont	24906	2.7	643077	1262	42.7	0.934
46	Virginia	110787	3.0	8631393	17753	44.3	0.880
47	Washington	184661	4.5	7705281	22377	38.4	0.902
48	West Virginia	62756	5.3	1793716	5674	45.2	0.866
49	Wisconsin	169635	3.0	5893718	17070	32.6	0.927
50	Wyoming	253335	4.1	576851	1258	12.9	0.911

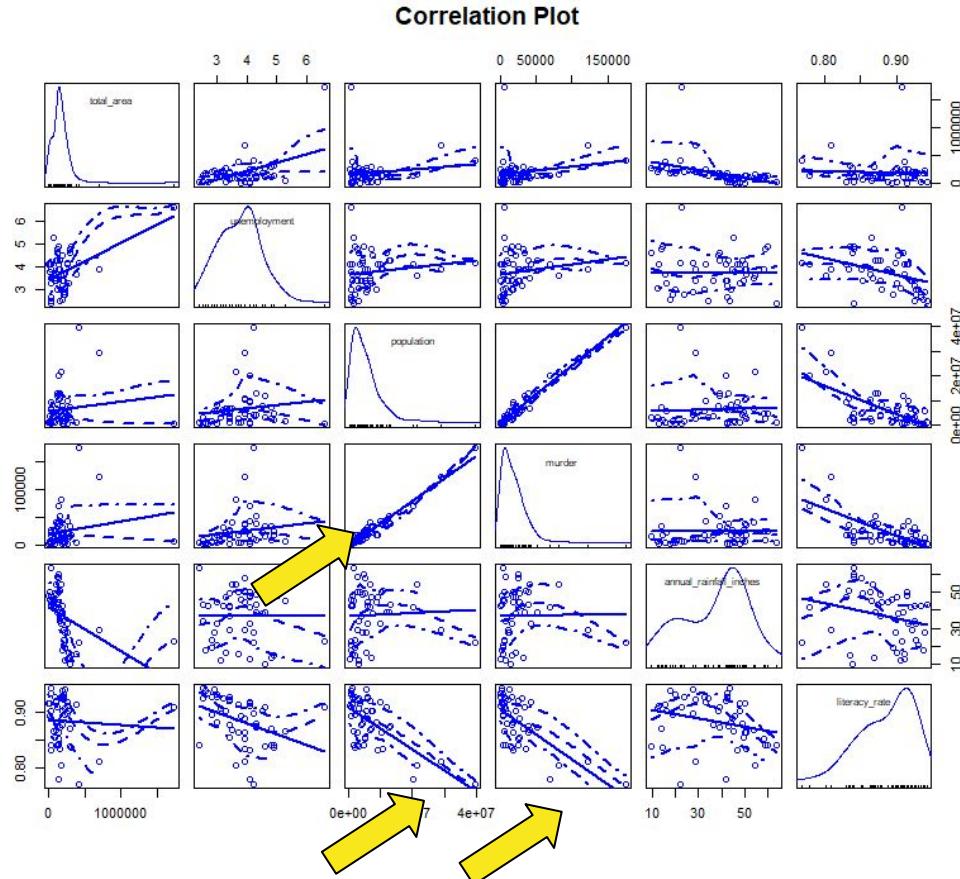
Console

Pearson correlation(2018 data)

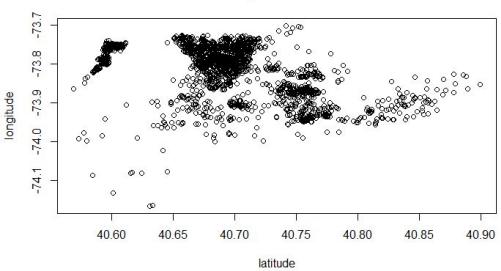
	total_area	unemployment	population	murder	annual_rainfall_inches	literacy_rate
total_area	1.00000000	0.4935022	0.12959572	0.18346301	-0.43092656	-0.05276924
unemployment	0.49350215	1.0000000	0.13965092	0.16540247	-0.01288670	-0.37208251
population	0.12959572	0.1396509	1.00000000	0.97584293	0.04786007	-0.65703598
murder	0.18346301	0.1654025	0.97584293	1.00000000	0.01289441	-0.66933022
annual_rainfall_inches	-0.43092656	-0.0128867	0.04786007	0.01289441	1.00000000	-0.25193921
literacy_rate	-0.05276924	-0.3720825	-0.65703598	-0.66933022	-0.25193921	1.00000000



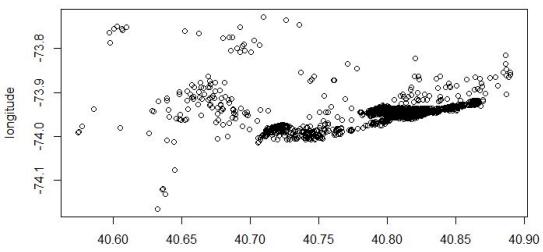
Pearson correlation plot on various var (2018 data of 50 States)



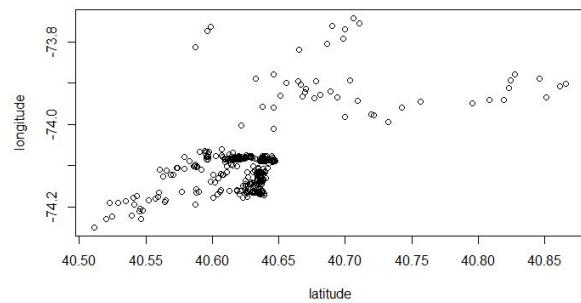
Queens Shooting Geometric Distribution



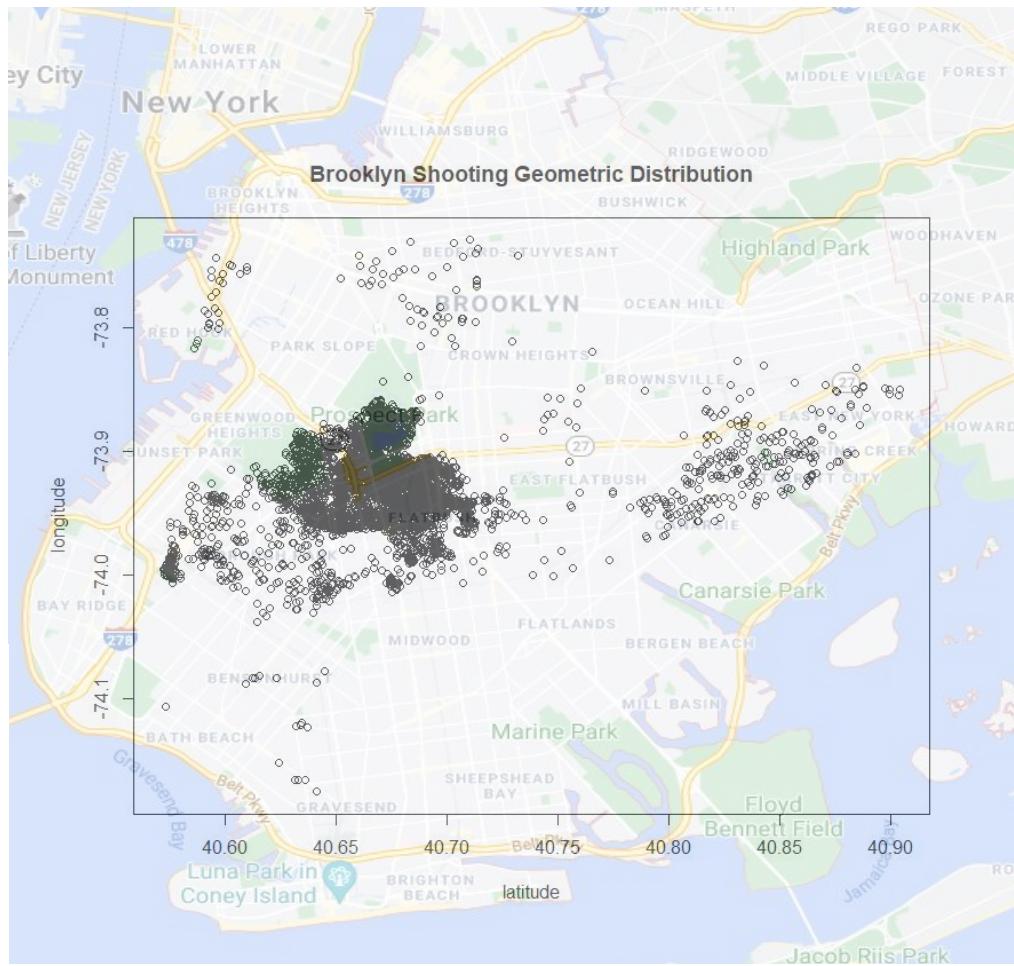
Manhattan Shooting Geometric Distribution



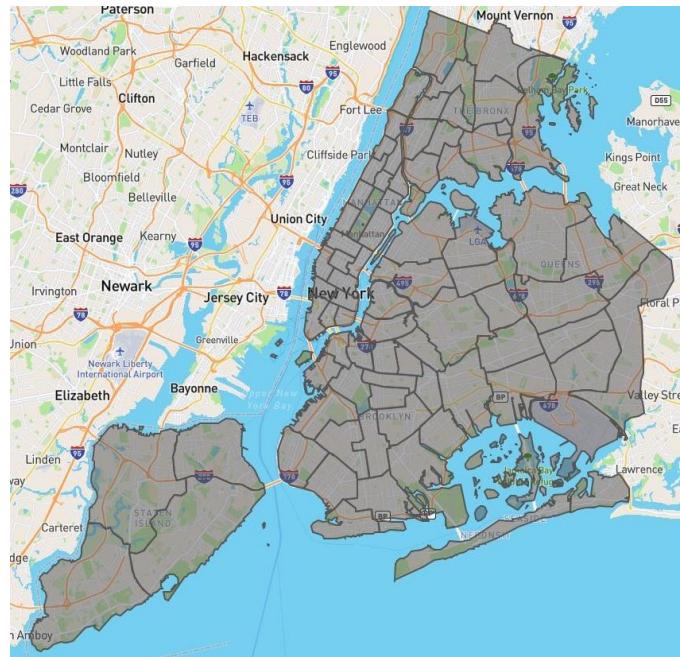
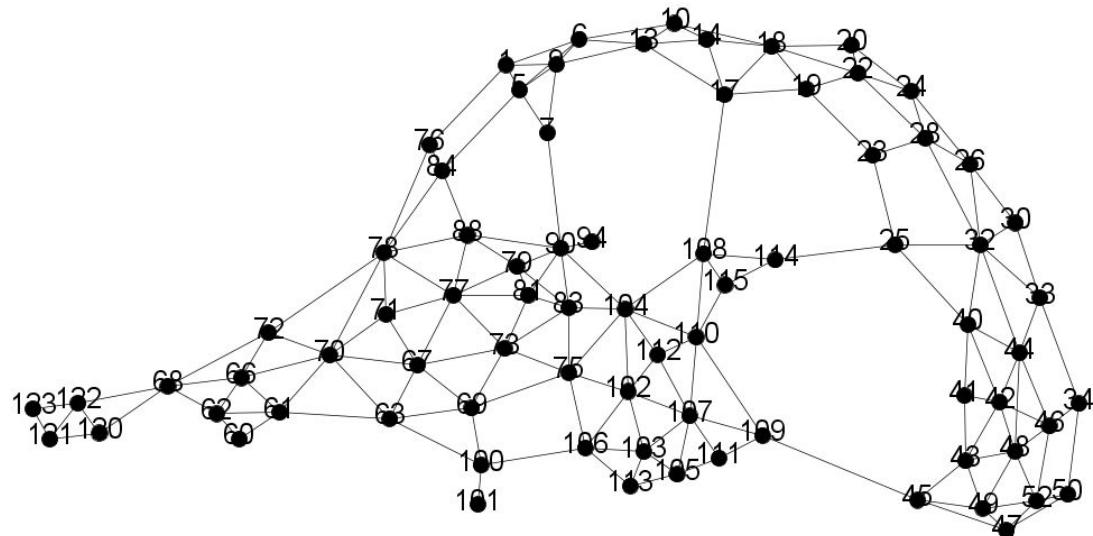
Staten Island Shooting Geometric Distribution



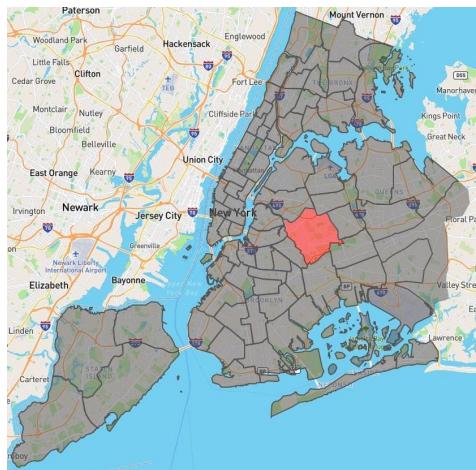
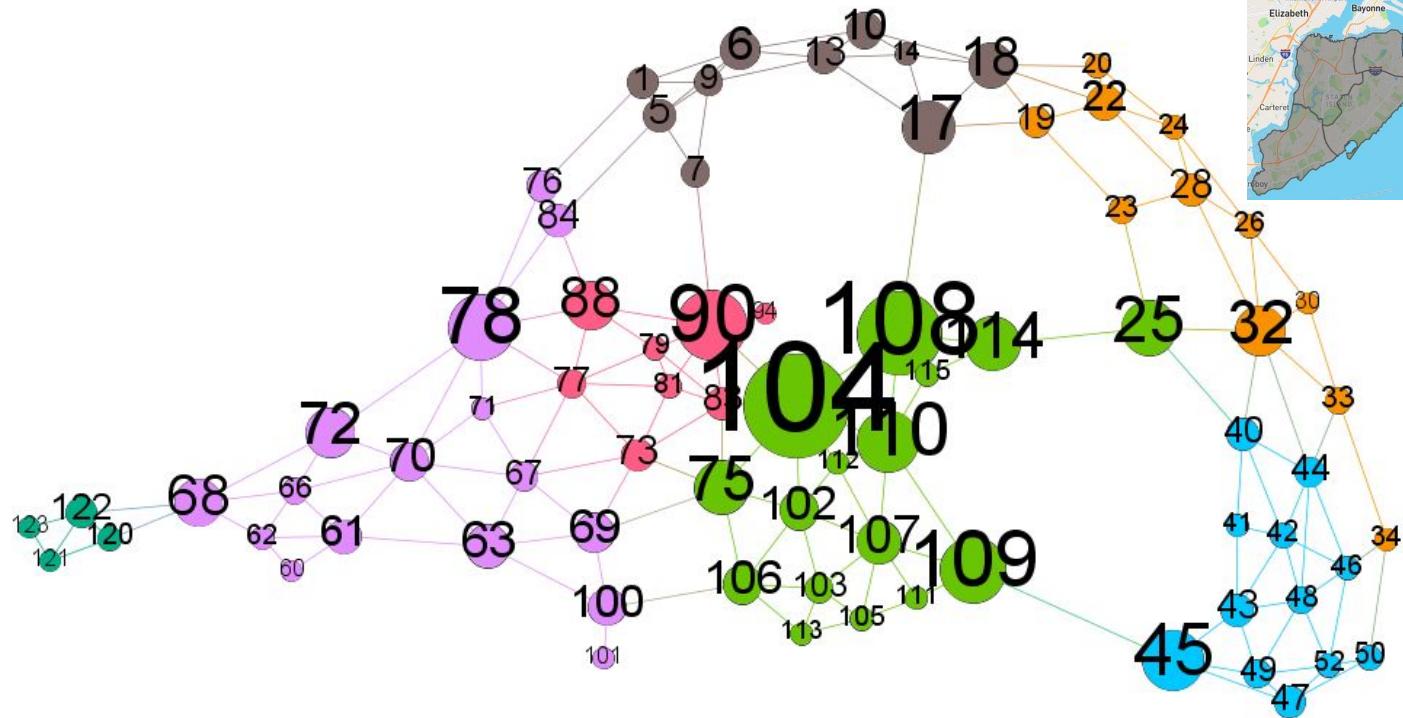
Brooklyn Shooting Geometric Distribution



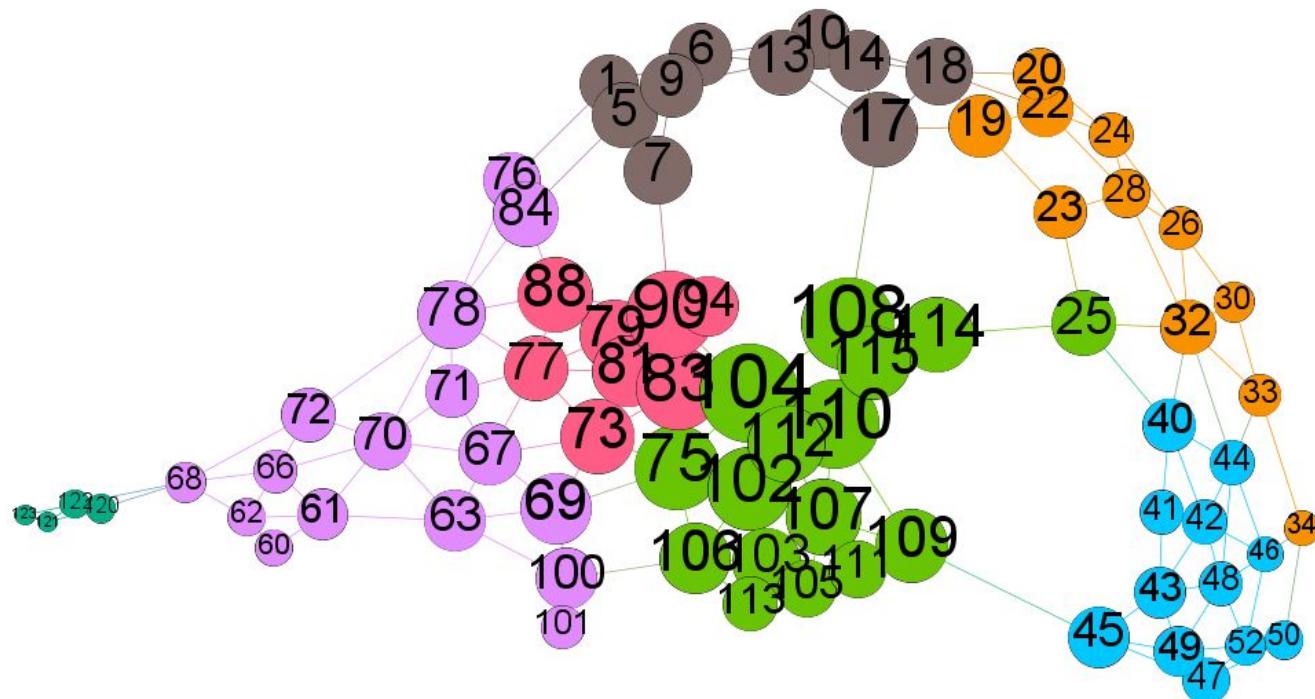
Network analysis - New York Police Precincts



Betweenness - New York Police Precincts



Closeness - New York Police Precincts



Network analysis - New York Police Precincts



Police Department
City of New York



Bill de Blasio
Mayor

Volume 28 Number 28

CompStat

Report Covering the Week
7/12/2021 Through 7/18/2021

Crime Complaints

Dermot Shea
Police Commissioner

104th Precinct

Volume 28 Number 28

CompStat

Report Covering the Week
7/12/2021 Through 7/18/2021

Crime Complaints

	Week to Date			28 Day			Year to Date*			2 Year			11 Year			28 Year		
	2021	2020	% Chg	2021	2020	% Chg	2021	2020	% Chg	2021	2020	% Chg	2021	2020	% Chg	2021	2020	% Chg
Murder	0	0	***.*	0	0	***.*	1	0	***.*	-50.0	0.0	-90.0						
Rape	1	1	0.0	1	2	-50.0	10	9	11.1	-33.3	11.1	-41.2						
Robbery	5	6	-16.7	12	16	-25.0	63	99	-36.4	-29.2	-53.3	-87.8						
Fel. Assault	5	8	-37.5	22	29	-24.1	130	167	-22.2	36.8	3.2	-19.8						
Burglary	5	5	0.0	16	20	-20.0	127	164	-22.6	22.1	-49.0	-88.6						
Gr. Larceny	12	8	50.0	38	34	11.8	252	309	-18.4	-26.3	12.0	-24.8						
G.L.A.	2	6	-66.7	13	19	-31.6	111	73	52.1	30.6	-44.5	-94.3						
TOTAL	30	34	-11.76	102	120	-15.00	694	821	-15.47	-5.19	-26.56	-83.16						

Volume 28 Number 28

CompStat

Report Covering the Week
7/12/2021 Through 7/18/2021

Crime Complaints

	Week to Date			28 Day			Year to Date*			2 Year			11 Year			28 Year		
	2021	2020	% Chg	2021	2020	% Chg	2021	2020	% Chg	2021	2020	% Chg	2021	2020	% Chg	2021	2020	% Chg
Murder	0	0	***.*	1	0	***.*	-50.0	0.0	-90.0									
Rape	0	1	-100.0	0	3	-100.0				8	9	-11.1	-38.5	-27.3	-82.2			
Robbery	5	6	-16.7	20	16	25.0	108	88	22.7	108	88	22.7	3.8	-26.5	-86.7			
Fel. Assault	8	3	166.7	20	27	-25.9	148	137	8.0	148	137	8.0	-9.2	27.6	-72.0			
Burglary	1	6	-83.3	7	12	-41.7	58	104	-44.2	58	104	-44.2	-9.4	-38.3	-94.9			
Gr. Larceny	6	7	-14.3	47	29	62.1	309	210	47.1	309	210	47.1	5.1	72.6	-53.5			
G.L.A.	4	1	300.0	9	6	50.0	97	54	79.6	97	54	79.6	234.5	131.0	-91.4			
TOTAL	24	24	0.00	104	93	11.83	735	604	21.69	9.70	24.37	-83.17						

77th Precinct

34th Precinct

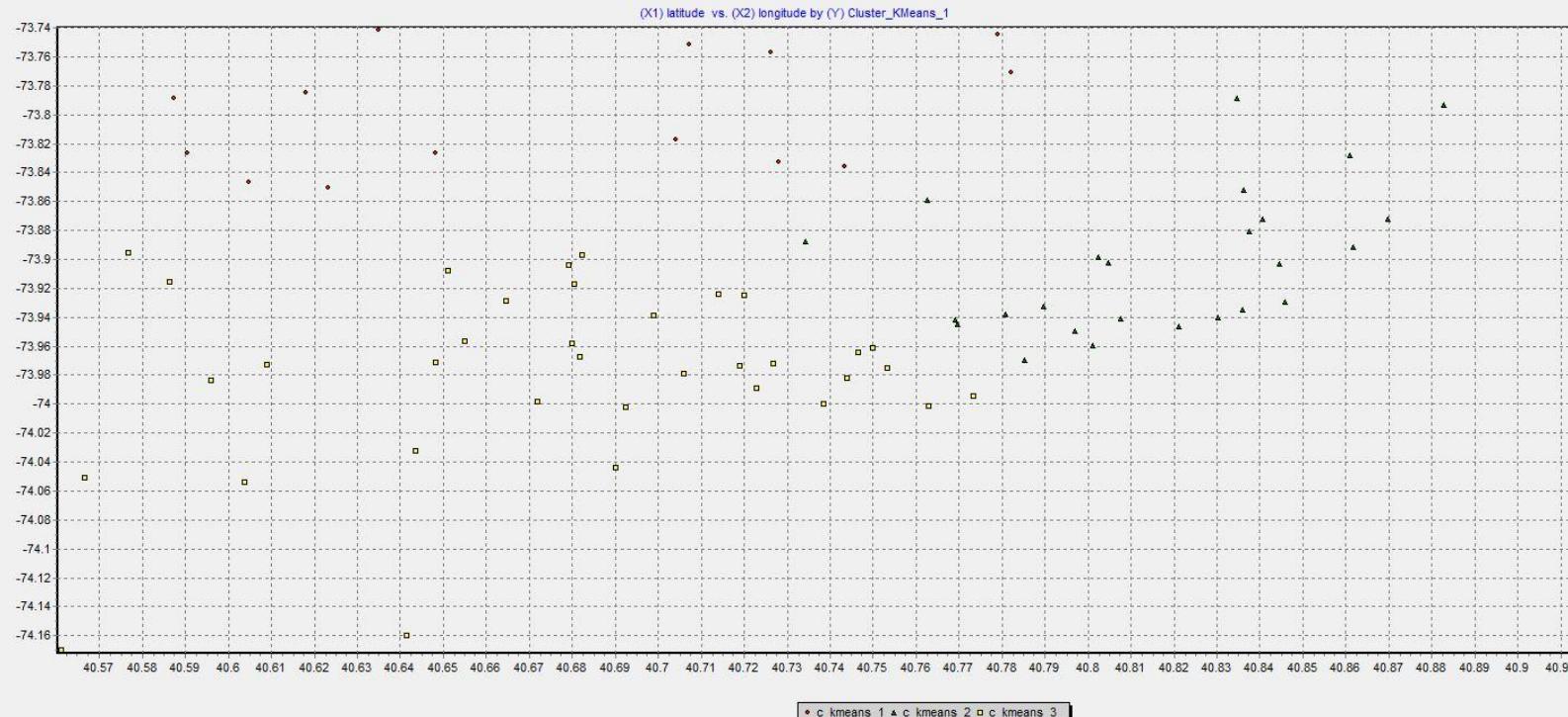
Cluster centroids

Attribute	Cluster n?	Cluster n?	Cluster n?
longitude	-73.798134	-73.903217	-73.978882
latitude	40.676956	40.820956	40.678958

K-means Clustering - New York Police Precincts

longitude

Cluster_KMeans_1



latitude

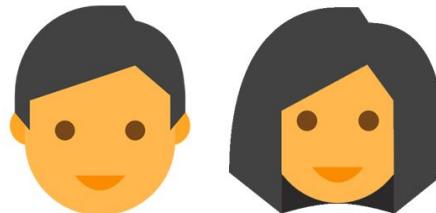
Possible areas of exploration

Are some inter-city precincts more important than others?



Centrality & Closeness

Are some age groups/genders targeted more often for crimes?



K-means clustering

What is the transportation monetary expenditure incurred?



Cliques and k-cores

Conclusion

Factors such as demographics, situational factors, crime type, and time impact crime rates.

Crime rates may be influenced by nation-wide events.

