Final Presentation

The Dapper Squirrels

Liqian Ma
Wentao Yao
Xingbang Liu

Northwestern | McCORMICK SCHOOL OF ENGINEERING

Theme

Misconduct analysis in terms of different locations and communities can be valuable. Is there over-policing in low socio-eco status neighborhoods? We could compare the low-income area data with high income area data. The income of the neighbor could be a factor to influence the "victim" narrative (complaint report). We plan to dive deep into the relationship between location, income level, and police misconduct.

Checkpoint 1

- What are the TOP 5 richest and lowest income neighborhoods?
  - **Riverdale Fuller Park Englewood East Garfield Park Washington Park**
  - **Forest Glen Lincoln Park Loop North Center Beverly**
- What are the neighborhoods' income and CRs(complaint record) per capita?
- What is the TRRS(tactical response report) per capita?
  - **N/A - can't associate income with beat areas**

Checkpoint 1

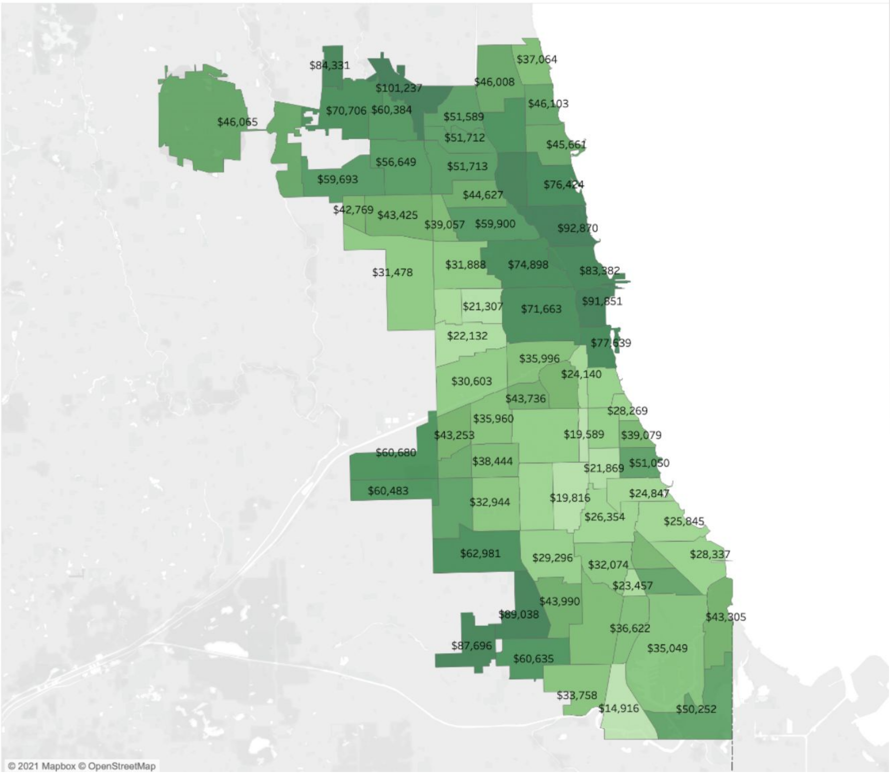
- What is the percentage of each race in the community?
  - **From *data_area* and *data_racepopulation***
- What are the top 5 streets in allegation counts for each beat area?

id	name	race	ratio
435	Roger Park	Asian	0.0643542516819234б
435	Roger Park	Black	0.2448439395610455
435	Roger Park	Hispanic	0.2414433292893643
435	Roger Park	White	0.4185324068968052б
435	Roger Park	Other	0.03082607257086136

beat_id	add2	allegation_count	rank
6	N WESTERN AVE	26	1
6	W IRVING PARK RD	9	3
6	W ADDISON ST	7	4
6	W BERTEAU AVE	5	5
6	N WESTERN AV	5	5
6	North CLARK ST	5	5

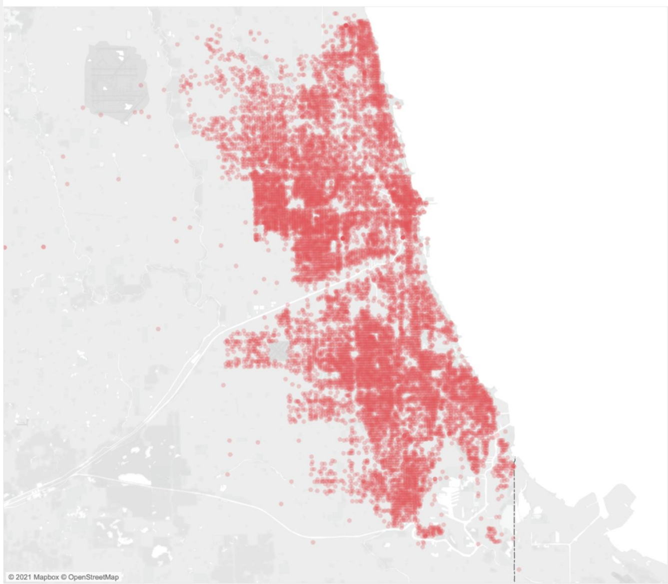
Checkpoint 2

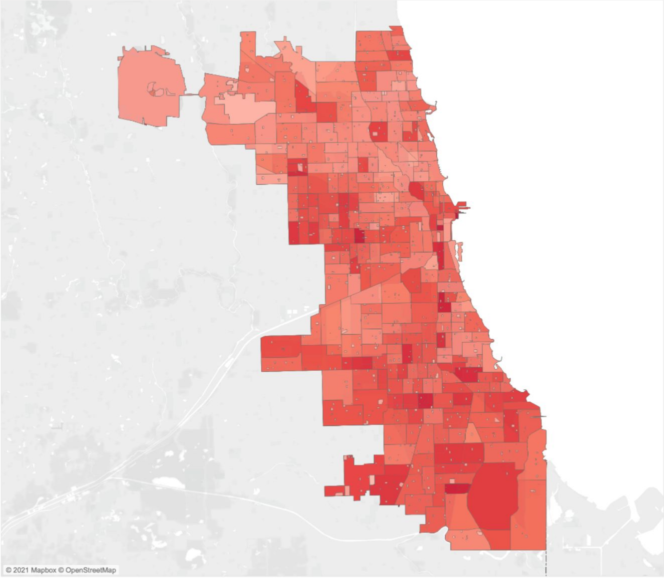
- Heatmap of the Income in different community



Checkpoint 2

- Scatterplot of Complaint Report per capita V.S. Tactical Response Report per capita. We could also consider lawsuits between “victims” and police officers; search warrants granted in each complaint?



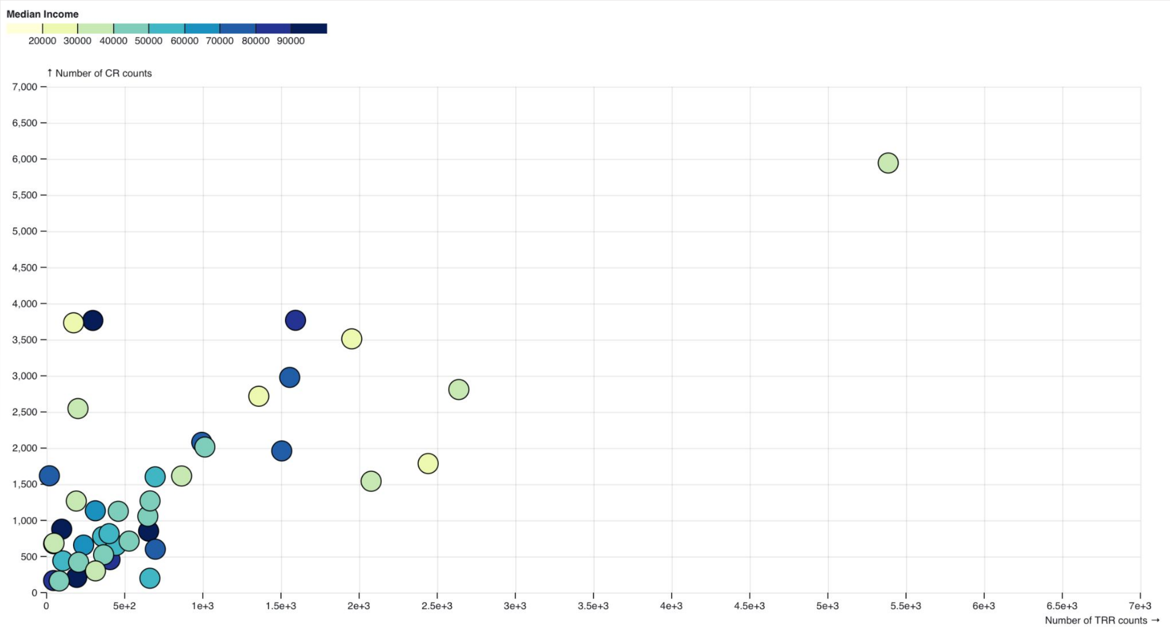


Beat_id to Community

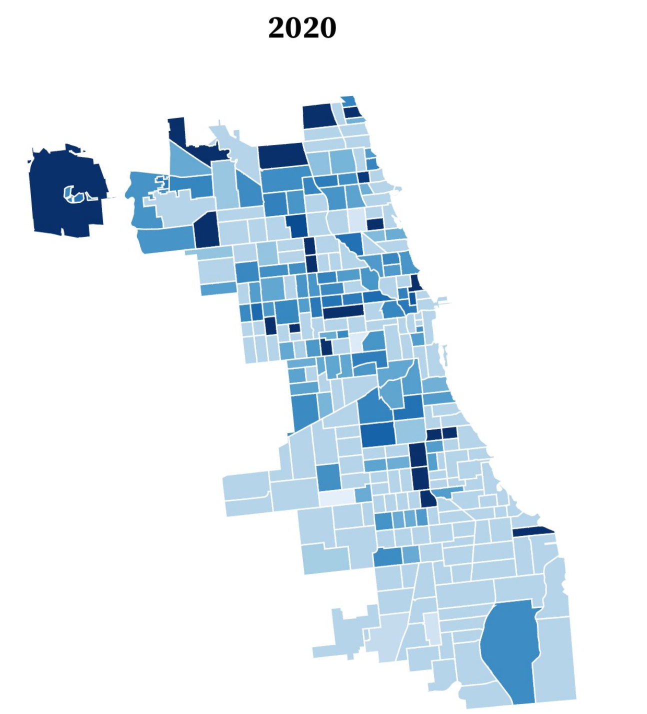
```
SELECT DISTINCT ON(1) table1.id as beat_id, table2.id as community_id
FROM (SELECT * FROM  data_area WHERE data_area.area_type ='beat')table1,
          (SELECT * FROM  data_area WHERE data_area.area_type ='community')table2
WHERE ST_Contains(table2.polygon, table1.polygon)
            or st_intersects(table2.polygon, table1.polygon)
```

Checkpoint 3

Highlighting the high and low socio-economic status communities with different colors and plot TRRs on them. Set up a time slider to see how it changes over time. Using color code(heat map) of A&A (dara_officer assignment attendance) in different neighborhoods. Set up a time slider to see how it changes over time.



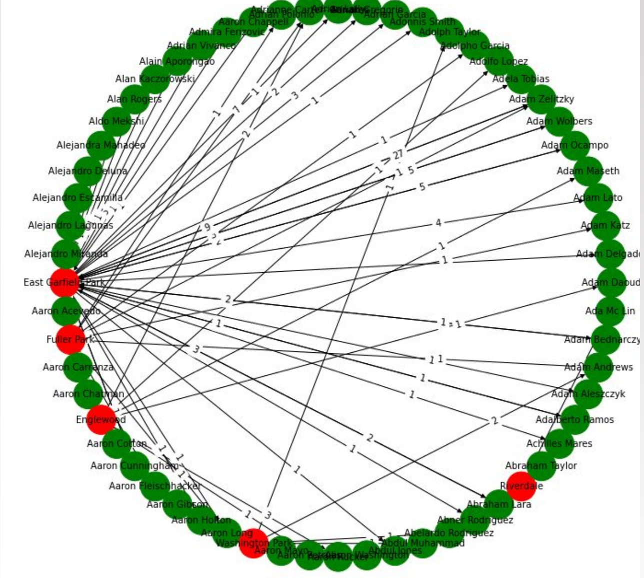
How CRs TRRs changes over time in different income communities.



Beat 212: 50.88 % Attendance

Checkpoint 4

- Making nodes of officers and victims by their income, race, locations, and even unsupervised machine learning models to learn the cluster and see if there is a potential connection between officers and victims.
- Network dynamics of co-accused in each cohort can be interesting. The analytics can be done with the following:
  - Make use of Triangle Count Algorithms for each cohort.
  - Make use of the Page Rank Algorithm to find the most connected officer in all cohorts.
  - How many CRs that officers have and how many co-accused for each cohort.
  - Compare the top k largest cohort of police officers in high and low socio-economy status.



label	count
17372	8316
3744	1636
29511	1224
11980	652
28273	596
28838	450
32014	364
32068	323
32382	257
26622	257
13631	256
14106	243
32274	211
32041	207
6534	187
18915	186
23787	173
2981	162
21912	155
23033	115

only showing top 20 rows

src	dst	relationship
Austin	Alan Krok	CR
Englewood	Ruth Johnson	CR
Chicago Lawn	Michael Mayhew	CR
South Deering	Nora Collins	CR
Woodlawn	Tracy Quarles	CR
East Garfield Park	Gerard Murphy	CR
Near North Side	Jose Zuniga	CR
Near North Side	Frank Cool	CR
Norwood Park	Jeffrey Fronczak	CR
Garfield Ridge	George Mc Murray	CR
Near West Side	Debra Ippolito	CR
Lower West Side	Jack Dedore	CR
Pullman	Joseph Buss	CR
Belmont Cragin	Latonia Harris	CR
Lincoln Square	Gail Martin	CR
Austin	Marienne Perry	CR
Englewood	Marilyn Uldrych	CR
Auburn Gresham	Michael Devine	CR
Beverly	George Porter	CR
Ashburn	Nicola Zodo	CR

only showing top 20 rows

id	inDegree
Joe Parker	129
Jerome Finnigan	124
Edward May	114
Charles Toussas	114
David Brown	109
Kevin Osborn	108
Maurice Clayton	107
Glenn Evans	106
Adam Zelitzky	105
Jerome Turbyville	99
Robert Smith	98
James Grubbs	93
Robert Johnson	93
John Carney	88
Gregory Jackson	87
Tyrone Jenkins	87
Broderick Jones	87
Kevin Ryan	85
Eugene Bikulcius	85
Edward Howard	83

only showing top 20 rows

id	outDegree
Austin	10470
West Englewood	7979
Loop	7927
Near West Side	7411
Near North Side	7327
Auburn Gresham	6009
Humboldt Park	5760
North Lawndale	5503
Englewood	5360
West Town	5267
South Shore	4932
East Garfield Park	4900
New City	4891
Roseland	4763
Chicago Lawn	4741
Logan Square	4368
Lake View	4114
Greater Grand Cro...	4088
Uptown	3833
Woodlawn	3752

only showing top 20 rows

Checkpoint 5

- Compare the Bigram and Trigram in the top 5 highest and lowest socio-economy status communities.
- For similar(in contents) complaint summary texts, are they having similar socio-eco status to each other?
- Is there any bias in the complaint report? In other words, is the complaint report narrative different from the incident summary?

Conclusion

- Social-economic status is closely related to over-policing.
- Richer areas tend to file more complaints than poorer areas.
- Poorer areas tend to receive more tactical responses from police officers.
- The over-policing problem need more evidences to be supported. We find there are more law enforcement in low-income communities, but more data and work are needed to prove there is a logical correlation.

Future Work

- High rate of complaints in high-income areas needs to be verified by detecting the potential bias in complaint summaries.
- The legality of tactical response needs to be verified by detecting the severity of incident or is the crime rate matches the tactical response rate.

The Dapper Squirrels

Northwestern | McCORMICK SCHOOL OF ENGINEERING