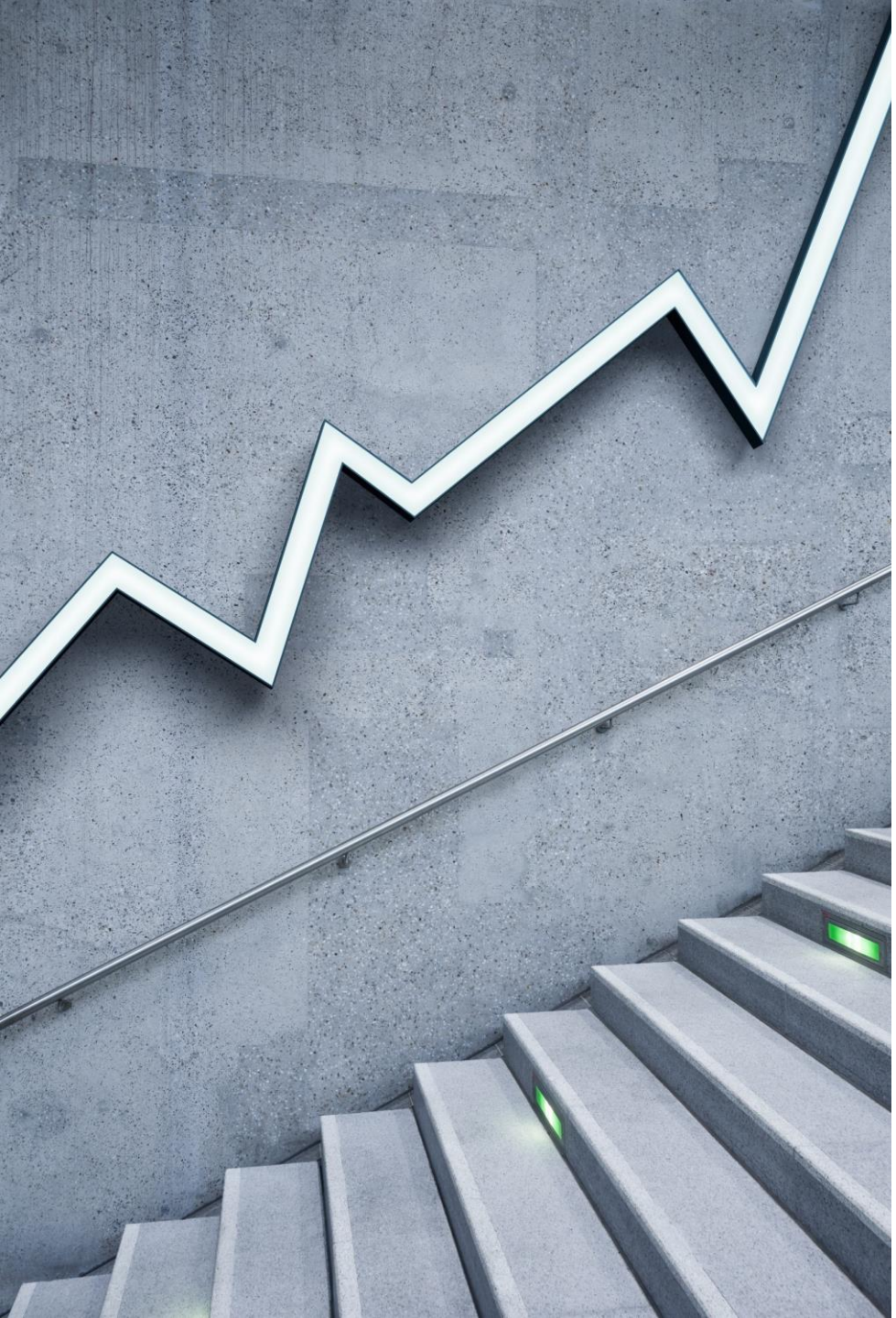


City of Philadelphia

CRIMES AND PUBLIC VENUES TYPES ANALYSIS

Content

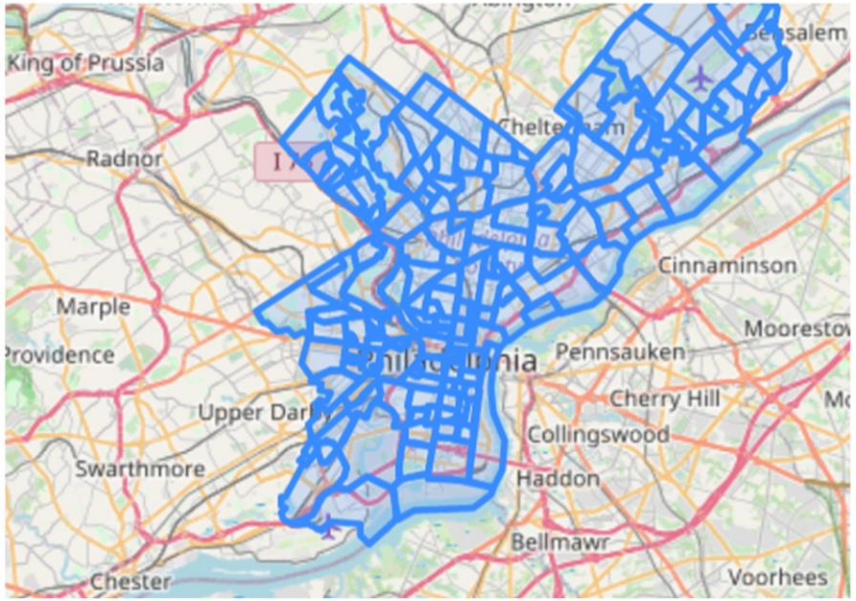
- Problem
- Data sources
- Results
- Conclusion

Problem

- Philadelphia is a city with diverse cultural and ethnical backgrounds

- Crime rate for serious crimes goes up[1]

- Can we see if there is relationship between neighborhood type based on venues and crime types?



[1]. https://www.inquirer.com/news/philadelphia-crime-levels-2019-shootings-homicides-police-20191230.html

Data sources:

- Philadelphia Police Department crimes data set. It includes both violent and nonviolent offences. It has location, general crime type, datetime and latitude and longitude of that address https://www.opendataphilly.org/dataset/crime-incidents
- Foursquare API – similar to the course material, used it for information about neighborhoods
- Neighborhood shape data in GeoJson format. This will be used to display neighborhoods borders on the map and also latitude and longitude of crime will be compared vs shapes from this source to map each crime into appropriate neighborhood. https://github.com/azavea/geo-data/tree/master/Neighborhoods_Philadelphia

Results

Out of 10 clusters I was able identify 2 that had a very distinct relationship between crime seriousness and venue types prevailing in that area:

- The neighborhoods where most frequent venues are “non bars/restaurants” but rather “non liquer/family interest” types like zoo/flea market/ flower shop/vintage store had less serious crimes prevailing in them.

- It doesn’t mean that these are generally safer without additional analysis as I haven’t considered number of crimes by type which should be done to provide a final validation. However this indicates that there is a possibility that liquor shops/bars tend to provoke people living in these neighborhoods to commit more serious crimes.

- For the rest of clusters it is hard to determine a crime type cluster prevailing as these are very mixed.

Cluster_venues	Cluster_Crime	
0	0	1.000000
1	7	0.225225
	6	0.207207
	8	0.180180
	0	0.153153
	4	0.081081
	1	0.063063
	5	0.054054
	9	0.018018
	2	0.009009
	3	0.009009
2	0	0.500000
	5	0.500000
3	7	1.000000
4	2	0.500000
	6	0.500000
5	6	0.500000
	8	0.500000
6	8	1.000000
7	5	1.000000
8	6	1.000000
9	7	0.323529
	6	0.264706
	4	0.117647
	8	0.117647
	5	0.088235
	3	0.058824
	0	0.029412

Name: Cluster_Crime, dtype: float64

Conclusion

The conclusions from the previous slide should be taken with a grain of salt as much more analysis is needed which is out of scope for the Coursera Capstone project.

However, even the first results show some potential in terms of showing trends between “serving liquer/non liquer” prevailing establishments and crime seriousness in that neighborhood.

The code included can be scaled and used for additional EDA (some tips and possible areas for improvement are included in the report)