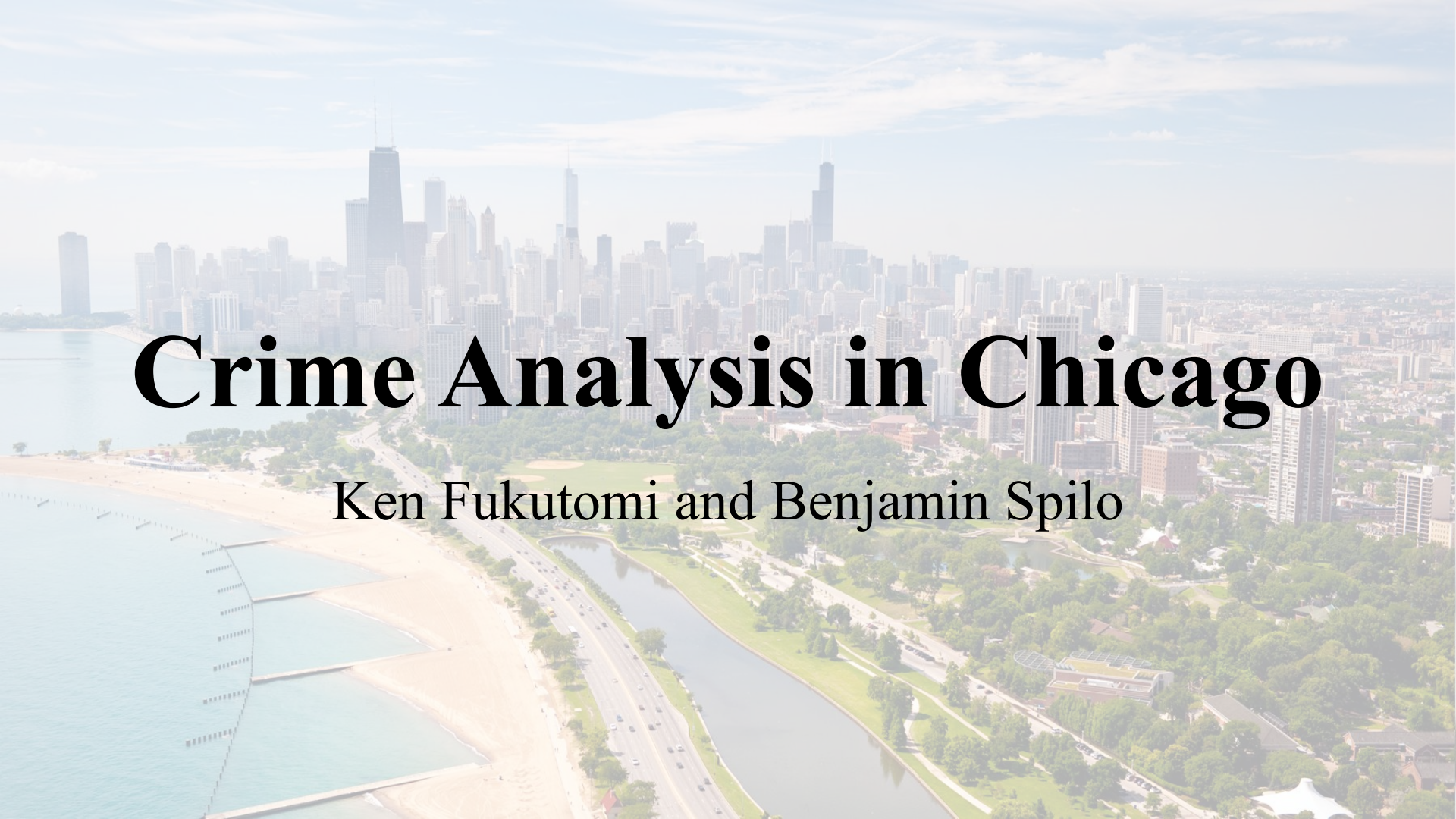

Crime Analysis in Chicago

Ken Fukutomi and Benjamin Spilo

Motivation

We are primarily motivated to work on this project to explore the city of Chicago and how traits of it's neighborhoods and regions affects its crime. We are interested in exploring the idea of predictive policing using historical data, and the ethical and practical concerns with its implementation. Through the analysis of the Chicago dataset, we are eager to uncover novel insights into how crime correlates with different social environments, thereby enhancing our overall understanding of the city's social fabric. By doing so, we aspire to contribute meaningfully to the discourse on urban development, safety, and data exploration in the future.

Idea

We're interested in how various physical attributes of neighborhoods and regions has affected crime in Chicago. We're looking at proximity to places of interest, housing prices, road networks, interest rate by quarter within the four year span, and other factors to see what could correlates to crime activity in Chicago. Additionally, given its an urban informatics class, we want to explore arrests that were driven by the recognition of systemic bias and the disproportionate policing of communities with lower incomes and minority populations. By doing these all together, we hope to better understand the city of Chicago itself.

Data

- Chicago Open Data Portal.
- US Census Income Data
- OSMnx and Pandana for POI data (OSM Geospatial Features)
- Zillow/Zestimate for housing price indices

```
for city in target:
    if not city in ["burnside", "riverdale", "mckinley park"]:
        cityname = f"{city}, chicago, illinois"
        crs = 3035
        graph = ox.graph_from_place(cityname, network_type='walk')
        graph = ox.projection.project_graph(graph, to_crs=crs)
        poi = ox.geometries.geometries_from_place(cityname, tags=tags)
        # project the place of interest
        poi = poi.to_crs(epsg=crs)

        max_time = 21
        walk_speed = 4
        for u, v, data in graph.edges(data=True):
            data['speed_kph'] = walk_speed
        graph = ox.add_edge_travel_times(graph)


        nodes = ox.graph_to_gdfs(graph, edges=False)[['x', 'y']]
        edges = ox.graph_to_gdfs(graph, nodes=False).reset_index()[['u', 'v', 'travel_time']]
```

Method

We plan to use geo-pandas for mapping. We also plan to use Pandas and Geo-pandas to do more traditional analysis. We also plan to do extra analysis with more complex statistical models, using many indices we will be collecting on the way.

Project Repo: https://github.com/kfukutom/urban-informatics-final/tree/main

Project Resources/Inspiration

- https://rpubs.com/hartwj/685827
- https://github.com/chris31415926535/walkscore (walk score api separate from POI calculator)
- https://towardsdatascience.com/the-math-behind-knn-3d34050efb71
- https://towardsdatascience.com/calculating-building-density-in-r-with-osm-data-e9d85c701e19