Introduction

- The United States is very segregated.[1]
- People of different income levels do not interact outside of their income group

Cosmopolitan mixing hypothesis:

"In large cities, the combination of increased population diversity, constrained space, and accessible public transportation will bring diverse individuals into close physical proximity with one another—reducing everyday socioeconomic segregation."

- Theory: Large cities have more interaction between people of different economic levels than smaller cities and towns.
- Is this true?

[1]: I mean, just look at it.
Why do we care?

- Segregation is bad\cite{citation needed}
- Specific reasons:
  - Reduces economic mobility
  - Creates health problems
  - Increases political polarization
- Reducing economic segregation will lead to improved public health and living conditions
Defining Exposure Segregation

- **Exposure Segregation (ES),** defined as the extent to which individuals of different economic statuses are exposed to one another, for each geographic area in the U.S.

- Higher ES == More segregated area
- Higher ES == People from different economic brackets don’t interact as much.

\[
\text{Exposure Segregation} = \text{Corr}(\text{SES}, \overline{\text{SES}}_{\text{exposures}}) = \frac{\text{cov}(\text{SES}, \overline{\text{SES}}_{\text{exposures}})}{\sigma_{\text{SES}} \sigma_{\overline{\text{SES}}_{\text{exposures}}}}
\]

(correlation between the socioeconomic status (SES) of individuals residing in that geographical area, and the mean SES those that they cross paths with.)
Measuring Economic Segregation

- Determine Economic Class:
  - Dox the users
  - Find their address during the nighttime
  - Look up value of that house address on Zillow

- Measure Economic Segregation:
  - Use GPS Pings to track location
  - One interaction counts as being within 50m for at least 5 minutes

Inferring home location

We first infer a user’s home latitude and longitude using their pings during local nighttime hours, based on best
Exposure Network

- Network of all interactions
  - Each node $n$ is one of $N = 9,567,559$ individuals
  - Each connection is one interaction
    - 1,570,782,460 interactions
    - Edge attributes: Time, Latitude, Longitude

- Total Locations: 3211
  - 382 cities
  - 2829 counties
Comparison with other methods

- Previous method of measuring segregation: Neighborhood Sorting Index
  - Assumes uniform and homophilic interactions
- However, most interactions occur outside of home tract
- New method is more robust

(a)

Conventional Segregation (exposure is uniform and only within home tract)  Exposure Segregation (real-world exposure)

One person  Exposure pairs (both / one / neither within home tract)
Results

- “Exposure segregation is extremely high in large cities.”
- “The top 10 largest MSAs, by population size, are 67% more segregated than small MSAs with less than 100,000 residents”
Why?

Gwin
Net Worth: $36.21

Daddy Bezos
Net Worth: $169,000,000,000
Small Town ($n \leq 100,000$)
Small Town (n <= 100,000)
Big Town ($n > 1,000,000$)
Supporting Data

- Higher population correlates
  - Higher variety of restaurants
  - Higher variety of prices per restaurant
  - Higher variety of house prices
  - Higher exposure segregation
Supporting Data

- Applies to all public places, not just restaurants
- Varies among the specific type of exposure hub
Alternate Explanation 1

- Constant Homophily Explanation
  - People naturally interact more with people with same economic status
  - Cities have more people of the same economic status
- Test: null network model, in which we preserve network nodes (individuals and their SES values) but randomize edges
  - Probability of exposure \( p_{i,j} \) between two individuals of \( ES_i \) and \( ES_j \)

\[
p_{i,j} \propto Similarity(ES_i, ES_j) = 1 - \frac{|ES_i - ES_j|}{\max(ES) - \min(ES)}
\]
Alternate Explanation 2

- Between Activity Theory
  - Large cities individuals choose different categories of activities which results in segregation
- Configuration model to configure network edges for each leisure category separately
- Maintain node degree, but filter by activity type
- No positive association between exposure segregation and population size

Exposure Segregation (Between-Activity Homophily)
Breakdown of Different POI Segregation

- Various types of POI have different exposure segregation
- More locations == More segregation
- Farther distance == Less Segregation
  - service small socioeconomically homogeneous communities
Golf Courses are Expensive

- Golf Courses are outliers (high distance, high economic segregation)
- Range of prices is far greater than any other industry
- Most expensive golf course 11717 times more expensive than cheapest
- Michelin star restaurant only 63 times more expensive than big mac

<table>
<thead>
<tr>
<th>Metropolitan Statistical Area (MSA)</th>
<th>POI Name</th>
<th>SES</th>
<th>Minimum Cost of Entrance ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicago-Naperville-Elgin, IL-IN-WI</td>
<td>Wynstone Golf Club</td>
<td>High</td>
<td>40,000</td>
</tr>
<tr>
<td></td>
<td>Flagg Creek Golf Course</td>
<td>Low</td>
<td>19</td>
</tr>
<tr>
<td>Los Angeles-Long Beach-Anaheim, CA</td>
<td>Coto de Caza Golf &amp; Racquet Club</td>
<td>High</td>
<td>45,000</td>
</tr>
<tr>
<td></td>
<td>Rancho Vista Golf Club</td>
<td>Low</td>
<td>129</td>
</tr>
<tr>
<td>Miami-Fort Lauderdale-West Palm Beach, FL</td>
<td>Old Palm Golf Club</td>
<td>High</td>
<td>199,200</td>
</tr>
<tr>
<td></td>
<td>Eco Golf Club</td>
<td>Low</td>
<td>17</td>
</tr>
<tr>
<td>New York-Newark-Jersey City, NY-NJ-PA</td>
<td>Scarsdale Golf Club</td>
<td>High</td>
<td>8,900</td>
</tr>
<tr>
<td></td>
<td>Weequahic Park Golf Course</td>
<td>Low</td>
<td>35</td>
</tr>
<tr>
<td>Phoenix-Mesa-Scottsdale, AZ</td>
<td>The Estancia Club</td>
<td>High</td>
<td>150,000</td>
</tr>
<tr>
<td></td>
<td>Peoria Pines</td>
<td>Low</td>
<td>28</td>
</tr>
</tbody>
</table>

= 63x

= 11717x
Solutions?

- What are some ways to reduce economic segregation?
Bridging Index

- Place frequently-visited POIs to act as “bridges” between diverse neighborhoods
  - Allows residents of nearby high-SES and low-SES neighborhoods to easily visit and cross paths
- **Bridging Index**: Measurement of how good of a “bridge” a place is

\[
\text{Bridging Index (BI)} = \frac{\text{Within Hub Economic Diversity}}{\text{Overall Economic Diversity}} = \frac{\sum_{i=1}^{K} |C_i| \cdot \text{Gini Index}(C_i)}{|V_A| \cdot \text{Gini Index}(V_A)}
\]

- Top 10 cities with highest BI 53.1% less segregated than 10 cities with lowest BI.
Bridging Index

- POI features that contribute to a high bridging index include
  - Low amount of these POI
  - Placement in between different economic area
**Bridging Index**

- Exposure segregation is negatively correlated with bridging index

![Diagram showing the relationship between bridging index and conventional segregation](image)
Does this work?

- Exposure Network predicts actual Facebook Social Connectedness
- Facebook Social Connectedness Index measures the relative probability of a Facebook friendship link between a given Facebook user in location i and a given user in location j
- Model shows that more connected (less segregated) communities have more friendships
- Strong correlations suggest that exposure segregation is likely related to segregation of friendships and other strong social ties.
Questions and Discussion