Evaluating Methods used to Quantify Racial Segregation

Abstract

Racial segregation has long been a problem in communities across the country. One approach to help understand such an important issue is to attempt to describe it quantitatively. Many metrics have been developed, all with various strengths and weaknesses, but none fully capture the nuances of this complicated issue. This work provides an overview of four of the mathematical approaches that have been developed to study segregation, explains how they function using small examples, and compares and contrasts their effectiveness in various situations. We then focus on segregation in Los Angeles (LA) County, including a detailed exploration of the most recent score proposed by authors Sousa and Nicosia, which conducts a random walk and outputs the number of steps it takes to reach all racial classes in the system. While we found there is a difference between the average step lengths of LA County vs. an unbiased null model, attempts to standardize outputs erases crucial data, and compressing this issue into one score is not representative of its complexity. This suggests that future exploration should attempt to study segregation more comprehensively rather than distilling an incredibly complicated and important issue into a single statistic. More work is needed to quantitatively represent the complexities of racial segregation in an effective matter.

1 Introduction

With a racially charged foundation and strong systemic discrimination, the United States has various levels of segregation in place throughout the country: The patterns of today's residential segregation closely follow the discriminatory practices developed in response to early racist ideals against marginalized groups [18]. The details in specific regions differ across the country, but in this work we focus on Los Angeles County. With over ten million citizens, LA county is the largest county by population in the United States of America [3], but with its large community comes major instances of racial inequality. For example, Black citizens represent only nine percent of the general population in LA County, yet comprise forty percent of the population experiencing homelessness [1]. This outcome, along with countless others that affect those in high-risk communities, can be traced back to the redlining practices appointed in the early establishment of the county – though they were implemented nearly a century ago their effects persist and harm racial minorities to this day, continuously contributing to instances of racial segregation.

Many attempts have been made to mathematically quantify the existence and severity of racial segregation in the U.S. Having a quantitative way to measure segregation allows us to more effectively contextualize the level of impact that racial discrimination has on communities, so that we can better understand and communicate the scale and scope of this critical problem. The ability to compare levels of segregation across different states and regions on a universal scale can help answer difficult questions about the allocation of scarce

resources, and comparing levels of segregation over time can help assess whether interventions have been effective. However, as of now there is no agreement between researchers on any single quantitative segregation measure to implement uniformly.

We will survey four measures of segregation that have appeared in the academic literature. Each captures certain aspects of segregation well, but fails to effectively describe other features. First, we cover a score that is widely used in the geography literature, Moran's I, which given a region divided into some geographic sub-regions considers the difference in the racial makeup of adjacent sub-regions [13]. Next, we examine the Dissimilarity Index: True to its name, it measures how closely the racial makeup of a sub-region aligns with the region's overall demographics [7]. We then move to the Clustering Propensity score, which looks at individuals rather than sub-regions and considers whether residents from a certain group tend to live next to those of the same group [2]. Finally, we consider a method that uses random walks: Given a graph with racial population proportions at each node, a random walk is conducted across nodes until all racial groups have been encountered, allowing us to evaluate the number of steps this takes [19].

We then consider what the random walk score tells us about Los Angeles County. We focus on this score because it provides a much more representative basis to analyze segregation, and effectively allows us to analyze multiple racial groups in an efficient manner. As discussed, analyzing the resulting levels of disparity from an objective, numerical point of view can also reveal characteristics about the overall structure of the region – however, we found that attempting to distill such salient data down to one score results in a substantial loss of context and information.

Our principal takeaways are that these scores on their own are insufficient, as none fully encapsulate the complete issue of segregation. We will see that some scores capture certain aspects of segregation better than others, but assert that using multiple scores can give a more complete picture than a single one on its own. Trying to distill a complicated issue down into a single score will never fully capture the complexity of this problem, and there is still work to be done in terms of finding a good combination of scores that can comprehensively describe the most relevant aspects of segregation.

This paper is organized as follows: We will first address the severely charged racial history of the U.S. and its impact on the development of Los Angeles County. An evaluation of previous measures of segregation will be discussed, along with the benefits and limitations of each measure. We provide several simple examples across all methods, followed by a full justification of the reasoning behind the method that was chosen for the rest of the analysis. We then provide a detailed explanation behind the measure and our adaptations, followed by an analysis of our results. We conclude with a discussion on limitations, further study, and a brief commentary on the implications behind this work.

2 Background

In order to comprehensively understand how racial segregation has been developed and maintained in specific regions, we will first explore the background and causes more broadly throughout the United States. This will allow us to recognize historical patterns of discrimination, form connections between the instances we see today, and develop a broader context about the impacts these have on citizens.

2.1 A (Brief) History of Segregation in the United States

Racial segregation can occur for many different reasons, often unintentional (a desire to form community and therefore residing close to family or others with similar racial backgrounds, for example). However, there have also been intentional, systemic discriminatory practices that enforce the separation of groups. This typically negatively affects those who belong in a racial minority.

It would be vastly inaccurate to disregard the history of slavery in the United States when discussing current practices of segregation: Though present informally for most of the U.S.'s history, segregation became systematic primarily following the enactment of the 13th Amendment [11] in order for those in positions of power to satisfy their desires to offset the recent ending of legalized slavery. Formerly enslaved people were denied the full rights and privileges of an average citizen so they would not constitute a threat to the previously upheld slave regime. This was not limited to law; these trends grew throughout housing, education, public accommodations, communities, businesses, and more, down to the subtle yet impactful details of separate doors, elevators, and drinking fountains [10]. Despite slavery being legally outlawed, several state segregation statutes came into being in order to uphold this notion of racial hierarchy. In analysis of early documentation during the period of legal enslavement, most laws contained explicit racial language, such as directly equating "slave" with African Americans, and using them interchangeably in written accounts [8]. The normalization of discriminatory language and practices ultimately contributed to instances of racial profiling that still exist to this day.

One of the most glaring representations of systemic segregation in minority communities occurred in the mid 1930s, resulting from the implementation of redlining. During this period the Home Owners' Loan Corporation (HOLC) collected a series of data that represented an area's likelihood of safe vs. risky mortgage security. Among the data was the neighborhood's quality of housing, the recent history of sale and rent values, and the racial and ethnic identity and class of residents [12]. The HOLC followed the guidelines of the federal housing administration's underwriting manual [9], which stated that "the infiltration of inharmonious racial groups will produce the same effects as those which follow the introduction of nonconforming land uses, which tend to lower the levels of land values and lessen the desirability of residential areas." To imply in a government document that "inharmonious" racial groups have a negative effect on housing values was detrimental to the establishment of integrated residency patterns. As time went on, this implementation created a self-fulfilling prophecy, where more immigrants/residents of color resulted in a lower neighborhood grade. These lower grades were the backbone of redlining; without access to mortgages these communities were unable to own property and build capital. These discriminatory ideas and practices ultimately led to the racial segregation that is still present today, and form the basis of the work conducted in this study.

2.2 Segregation Practices in Los Angeles County

Now that we have a stronger foundation on which to analyze racial segregation, we can apply this to particular regions. Specifically, we will examine these effects on Los Angeles, CA. A bustling city with boundless opportunities, LA has historically drawn citizens with diverse interests and backgrounds. As of July 2021, the total population of LA County reached 9,829,544 people: 49.1% Hispanic or Latino, 25.3% White alone (not Hispanic or Latino), 15.6% Asian, 9.0% Black or African American, 1.5% American Indian and Alaskan Native, and 0.4% Native Hawaiian and other Pacific Islander (with note that 3.3%

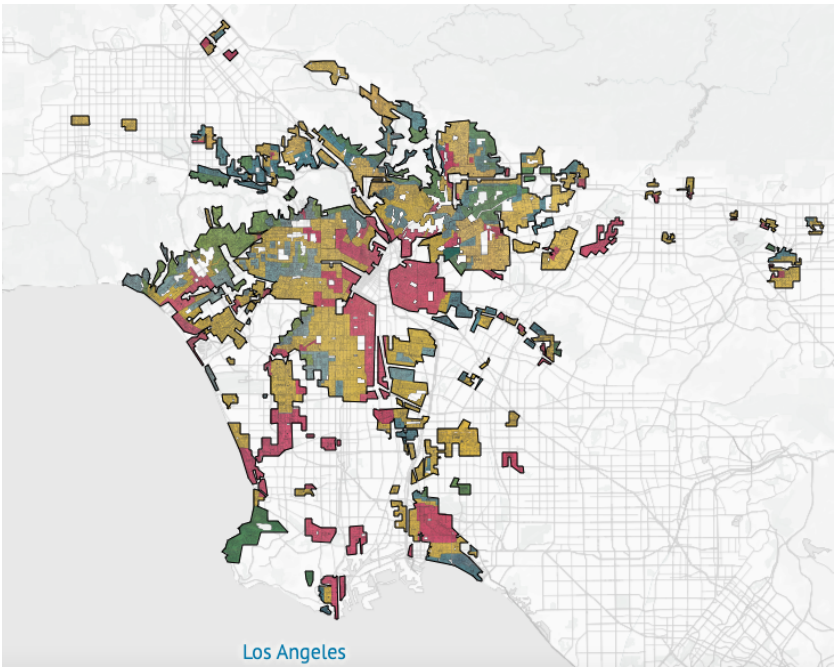


Figure 1: Original redlining map of Los Angeles, CA, created by the HOLC [12]. Areas colored red were graded "hazardous," yellow "definitely declining," blue "still desirable," and green "best."

of those surveyed identify as two or more races) [21]. However, diversity does not equate to integration; like most other major U.S. cities, there exists a high degree of residential racial segregation [20] due to the practices mentioned in the previous section. As the HOLC crafted the security map of Los Angeles in 1939, both class hierarchies and racial segregation worked together to format the population distribution of this region of Southern California. We can see the original redlining patterns created by the HOLC in Figure 1.

This was perpetuated in early years primarily through the implementation of zoning and restrictive covenants (including racially restrictive covenants), and Contracts, Conveyances, and Restrictions (CCRs). The 1924 Code of Ethics Article 34 implicitly defined the ideal neighborhood as segregated by race and class: "A realtor should never be instrumental in introducing into a neighborhood a character of property or occupancy, members of any race or nationality, or any individuals whose presence will clearly be detrimental to property values in the neighborhood" [14]. Ultimately, these practices promoted the prohibition of certain racial or ethnic groups from either owning or occupying a property, until the Supreme Court later deemed them unlawful [15]. This impacted integrated development throughout the years, as the severe aforementioned restrictions on mortgages and loans took their toll on the racial diversity/homogeneity of the city. This resulted in minority populations densely concentrated in neighborhoods with histories of extreme disadvantage [4]. The effects of redlining are still evident in current residential trends: consider the racial distribution of LA County (as of 2020) seen in Figure 2, where the clustering patterns of racial minority groups closely follow those of the original HOLC guidelines. One example (shown in Figure 3) is that the most prominent "hazardous" region, the downtown Los Angeles business

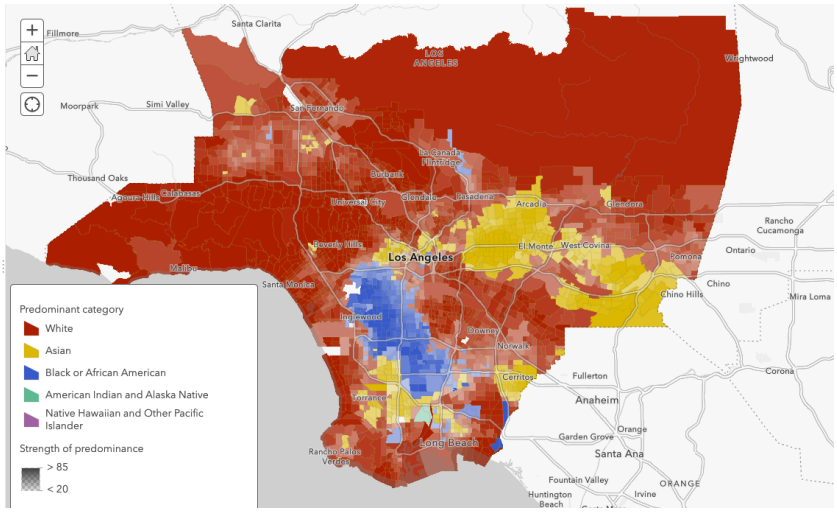


Figure 2: Census tracts within Los Angeles County (decennial redistricting data, 2020 (PL 94-171)) by race, created with ArcGIS. Note that the Hispanic/Latino category is considered by the U.S. Census to be an ethnicity rather than a race.

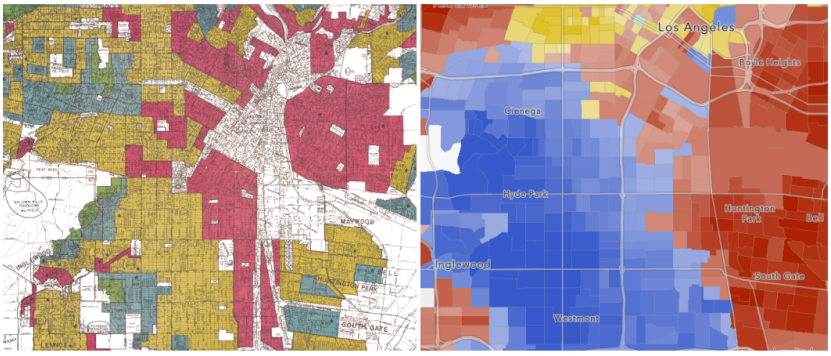


Figure 3: On the left: The Los Angeles business district, a subset of LA County deemed by the HOLC to be dangerously hazardous. On the right: Approximately the present-day configuration of the Los Angeles business district, a majority Black/African American region. Coloring in the left and right panels aligns with those in Figure 1 and Figure 2, respectively.

district, was located just east of present day Inglewood, a majority Black/African American region.

3 Techniques for Quantitatively Measuring Segregation

Although several measures to evaluate segregation have been explored, there is no consensus among experts on a single method to implement uniformly. Before we showcase these measures in action, we will provide some helpful definitions and several clarifying examples.

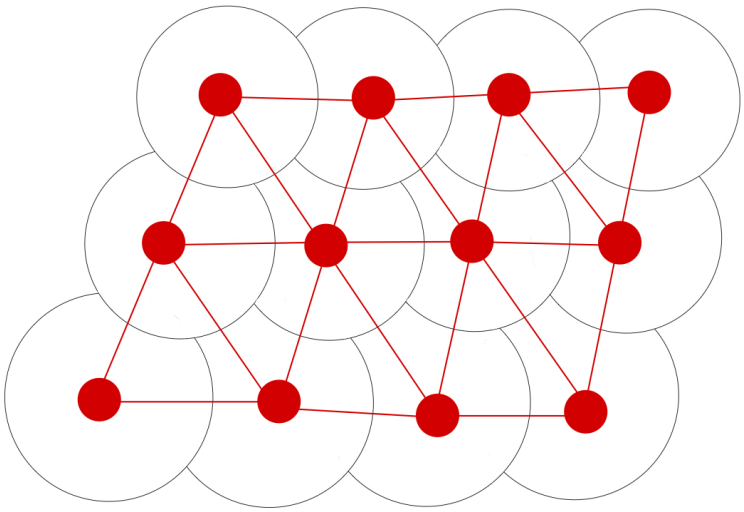
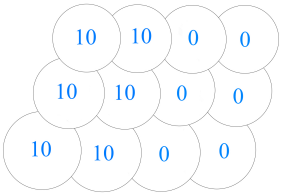


Figure 4: Dual graph of a 12-unit example as a 3×4 triangular lattice.

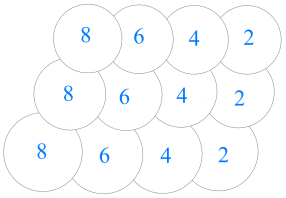
Through a comprehensive overview conducted by Rodriguez and Vorsatz [16], it was noted that the main mathematical devices developed to measure segregation are *segregation indices*: Formally defined as some function $S : N \to [0, 1]$ that maps a distribution N into the unit interval (where the maximum value of segregation occurs when each unit contains individuals from only a single group), standard indices output some value between 0 (minimal segregation) and 1 (maximum segregation) for a region of interest.

For a given region, a *dual graph* can be constructed by placing a node at each geographic unit (such as a census block) with an edge between two nodes if the corresponding units are adjacent. Additional information, such as population and demographics, can be added at each vertex v of a given region. This allows us to evaluate the structure of a region through the lens of graph theory.

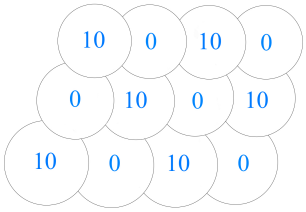
In order to properly understand the various mathematical measures of segregation we consider, we will be demonstrating the outputs of these scores on the following clarifying examples: Given that populations typically do not reside in a square grid, the dual graph we consider for each example is a 3×4 triangular lattice (from 12 underlying geographic units), shown in Figure 4, where the black lines denote the boundaries of the geographic units and the red nodes and edges comprise the dual graph. For the sake of simplicity, we will assume that there are 10 people living in each geographic unit. We will also assume that there are only two populations groups in the region, red and blue, and will compute all scores with x representing the jurisdiction-wide blue population. Visualizations of the different example population distributions we consider can be seen in Figure 5.



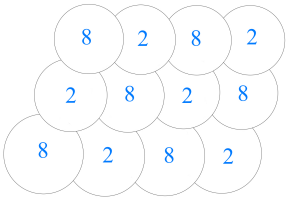
(a) A segregated region, with six connected units containing entirely blue residents and six connected units containing no blue residents.



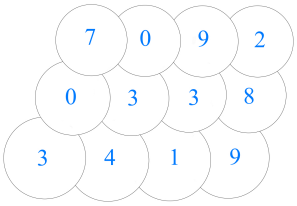
(b) A partially segregated region, with a gradual left-right decline of blue residents



(c) An integrated region (between units but not within units): six alternating units contain entirely blue residents and the remaining units contain no blue residents.



(d) A partially integrated region, with six alternating units containing eight blue residents and remaining units containing two blue residents.



(e) A randomly selected region, determined by generating a random number between 0-10 to determine the population of blue residents in each unit.

Figure 5: Simple examples for evaluating different segregation indices. For two population groups, blue numbers in each circle represent the unit-wide population of blue residents.

We believe that this set of examples provides a strong enough basis to evaluate the output of each measure – there is a wide enough span of population distributions to see several different examples of varying segregation/integration. Scores will be calculated using the methods described from each respective measure.

The measures we explore in the following sections are just a subset of those that have been studied throughout various fields of mathematics. For each measure we show the necessary calculations for a different example from Figure 5 in order to demonstrate this process and build intuition. We present all four segregation scores for all five examples, but omit the remaining calculations.

Table 1 displays the notation used throughout the next sections in the calculation of each measurement.

Variable	Definition
n	Total number of nodes in the dual graph
$\|E\|$	Total number of edges in the dual graph
$i \sim j$	Node i is adjacent to node j
$\bar{x}$	Total count of x population
x_i	Within-unit count of x population
$\bar{p}$	Count of total population
p_i	Within-unit count of total population
x_0	Average fraction of x population across all nodes

Table 1: Notation used throughout Section 3 in the calculation of each score.

3.1 Moran's I: P.A.P Moran (1950)

The first measure we will discuss is the standard spatial statistic used in geography literature: Moran's I. Developed by P.A.P Moran in 1950 [13], given numerical values (such as population) associated with the nodes of a dual graph, Moran's I returns a real number between -1 and 1 [6]. The interpretation is that values near 1 indicate extreme segregation (very clustered like-populations), values near zero indicate no significant pattern, and negative values flag "anti-segregation" (where people are more likely to live next to those of a different race). Moran's I essentially examines where values are positive or negative and analyzes any patterns in such areas.

While Moran's I can be defined with a more general formula [6], the definition for the rest of this analysis uses the within-unit count of x_i population rather than the fraction, as this approach has data for Los Angeles easily available. Let $x_0 = \frac{\bar{x}}{n}$ be the average presence of x population over all nodes. For n as the number of nodes and $|E|$ as the number of edges, Moran's I can be calculated as follows:

$$I = \frac{n}{|E|} \cdot \frac{\sum_{i \sim j}(x_i - x_0)(x_j - x_0)}{\sum_i (x_i - x_0)^2}$$

where $i \sim j$ if the two nodes are adjacent.

As written, the value of I could be anywhere between -1 and 1. Note that as each measure in this section will output a score between 0 and 1, in order to evaluate all results on an equal basis the remainder of this discussion scales Moran's I as such: $S_I = \frac{I+1}{2}$.

Researchers in the field of geography aimed to adapt these scores in order to account for potentially critical spatial relationships. However, a major concern with using Moran's I is that changing the aggregation level has a drastic impact on the output (frequently referred to as the Modifiable Areal Unit Problem). The result depends heavily on the choice of geographical units – the output will differ based on how large or small the chosen units are. This is a strong concern when attempting to relate scores of different regions that may be separated into varying units, and it is extremely unreliable to have an output rely on how a unit is defined.

3.1.1 Example Outputs of Moran's I

We will consider example 5b to demonstrate Moran's I: Here, the total population of x residents is $\bar{x} = 60$, and there are a total of 12 units in the region (with 23 total edges

between them). Therefore, the average presence of x residents across all nodes is $x_0 = 5$. There are only x population groups of eight, six, four, and two residents, so we can simplify the expression as follows:

$$\begin{aligned} I &= \frac{n}{|E|} \cdot \frac{\sum_{i\sim j}(x_i - x_0)(x_j - x_0)}{\sum_i (x_i - x_0)^2} \\ &= \frac{12}{23} \cdot \frac{5\left((8-5)(6-5) + (6-5)(4-5) + (4-5)(2-5)\right) + 2((8-5)^2 + (6-5)^2 + (4-5)^2 + (2-5)^2)}{3(8-5)^2 + 3(6-5)^2 + 3(4-5)^2 + 3(2-5)^2} \\ &= \frac{12}{23} \cdot \frac{65}{60} \\ &= \frac{13}{23}. \end{aligned}$$

Recall that we need to scale the output as such:

$$\begin{aligned} S_I &= \frac{I+1}{2} \\ &= \frac{\frac{13}{23}+1}{2} \\ &= \frac{18}{23} \\ &= \boxed{0.7826}. \end{aligned}$$

Outputs for Moran's I for all examples are shown in Table 2.

Example (5)	**Output**
Total Segregation (5a)	0.7826
Gradual Segregation (5b)	0.7826
Total Checkerboard (5c)	0.2609
Partial Checkerboard (5d)	0.2609
Random (5e)	0.5104

Table 2: Moran's I outputs for the examples introduced in Figure 5.

Notice that the values are equal for Total Segregation and Gradual Segregation, as well as for Total Checkerboard and Gradual Checkerboard. This is due to them simply being differently scaled versions of each other, demonstrating the Modifiable Areal Unit Problem discussed previously.

3.2 Dissimilarity Index: Duncan and Duncan (1955)

The next measure we will discuss is one of the most widely used segregation indices. The Dissimilarity Index [7] measures how closely subarea demographic proportions match the demographic proportions of the larger area – It measures "evenness," or the consistency of the levels of a sub-population over the units that make up a jurisdiction. This index was

developed shortly after Moran's I, and is arguably a much simpler approach to interpreting the problem at hand.

The idea of the Dissimilarity Index is, for each node, to look at how the node's population from one group (x_i) differs from what you would expect at that node if the group was evenly distributed across the region, which would be $p_i \cdot (\frac{\bar{x}}{\bar{p}})$. Adding the absolute values of these differences for each unit, rearranging terms, and scaling such that the result is always between 0 and 1 gives the following formula for the Dissimilarity Index:

$$D(x) = \frac{1}{2\bar{x}(\bar{p} - \bar{x})} \sum_i |x_i\bar{p} - p_i\bar{x}|.$$

But a severe problem with this method, as noted by Alvarez et.al [2], is that the Dissimilarity Index is given by summing over the nodes without reference to adjacency, so it does not take into account the *spatial relationship* between units. Thus this score equates neighboring units to those on opposite sides of the region of interest, potentially omitting crucial spatial information, as we will see in Section 3.2.1.

3.2.1 Example Outputs of the Dissimilarity Index

In order to showcase the Dissimilarity Index in action, consider example 5a: Here, the total population of residents is $\bar{p} = 120$, and the total population of x residents is $\bar{x} = 60$, and p_i will always be 10. The calculation can be completed as follows:

$$\begin{aligned} D(x) &= \frac{1}{2\bar{x}(\bar{p} - \bar{x})} \sum_i |x_i\bar{p} - p_i\bar{x}| \\ &= \frac{1}{2(60)(120 - 60)} \sum_i |120x_i - 10(60)| \\ &= \frac{1}{7200} \sum_i |120x_i - 600|. \end{aligned}$$

For a given node in this example, x_i will always be either 10 or 0:

$$\begin{aligned} &= \frac{1}{7200} \cdot 6(600) + 6(600) \\ &= \frac{1}{7200} \cdot 7200 \\ &= \boxed{1}. \end{aligned}$$

Outputs for the Dissimilarity Index for all examples are shown in Table 3.

Note that the output for Total Segregation and Total Checkerboard is the same: this is due to the fact that they both have the same population distribution, showcasing the drawback of this score's indifference to adjacency.

Example (5)	Output
Total Segregation (5a)	1
Gradual Segregation (5b)	0.4
Total Checkerboard (5c)	1
Partial Checkerboard (5d)	0.6
Random (5e)	0.5749

Table 3: Dissimilarity Index outputs for the examples introduced in Figure 5.

3.3 Clustering Propensity (CAPY): Alvarez et al. (2018)

The clustering propensity score, designed by Alvarez, Duchin, Meike, and Mueller [2], measures the clustering level of one or more subgroups within a population. This score focuses on how clustered a given region is, based on if people from a certain group tend to live next to those of the same group or those of a different group. This is done by taking the edges that connect x population to either x or y population and recording the share of xx edges. The argument is that CAPY scores successfully discern qualitatively important differences while providing a stabler baseline for interpretation than traditional segregation scores.

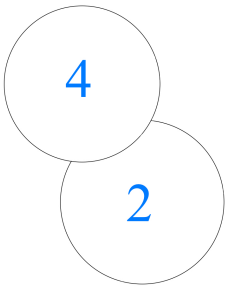
To measure the extent to which people of one demographic tend to live next to each other (rather than next to those of a different demographic), this measure considers an alternative to the dual graph, called an *exploded graph* (shown in Figure 6c). In the exploded graph, each node of the dual graph is replaced by a complete graph with a node for each individual. The clustering propensity scores are calculated using these exploded graphs, which allows for the analysis of both within-unit adjacencies and neighboring-unit adjacencies.

For distinct populations x and y, let x_i, y_i be the respective integer-valued populations for each unit. We can thus define the following:

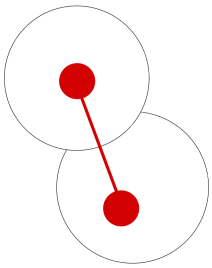
$$\langle \mathbf{x},\mathbf{y} \rangle = \sum_i x_i y_i + \sum_{i \sim j} x_i y_j + x_j y_i$$

The term $x_i y_i$ calculates how many edges there are (in the exploded graph) between people from group x and group y within node i, while the term $x_i y_j$ calculates how many edges there are (in the exploded graph) between people from group x at node i and people from group y at adjacent node j (and vice versa for the term $x_j y_i$). In all, $\langle \mathbf{x},\mathbf{y} \rangle$ calculates the total number of edges (in the exploded graph) between people in group x and people in group y. The number of edges between two people who are both in group x is given by $\frac{1}{2}\langle \mathbf{x},\mathbf{x} \rangle$: the edges between two different x-type vertices are over counted by a factor of two in $\langle \mathbf{x},\mathbf{x} \rangle$, which must be accounted for. Note this definition counts a vertex as being adjacent to itself; that is, if there are x_i nodes in the exploded complete graph for vertex i, it assumes there are $\frac{x_i^2}{2}$ adjacencies between x-type people within this node, rather than the more typical but less convenient count of $\binom{x_i}{2} = \frac{x_i(x_i-1)}{2}$.

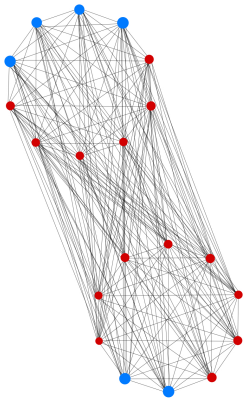
Using this notation, the CAPY Edge score looks at the edges incident to an x vertex (of which there are $\frac{1}{2}\langle \mathbf{x},\mathbf{x} \rangle + \langle \mathbf{x},\mathbf{y} \rangle$) and considers which fraction of them connect two x vertices (of which there are $\frac{1}{2}\langle \mathbf{x},\mathbf{x} \rangle$ such edges). The same is done for the group of y residents, and the results are averaged together. After simplifying, the CAPY Edge score can be written as:



(a) An example region to display an exploded graph, with 1 unit containing 4 blue residents and 1 unit containing 2 blue residents.



(b) The corresponding dual graph for the region shown in Figure 6a



(c) An example of an exploded graph on a region that represents a subsection of the partially segregated example shown in 5b. Note that each region transforms into a complete graph, and the number of edges significantly increases once exploded.

$$Edge(\mathbf{x},\mathbf{y}) = \frac{1}{2}\left(\frac{\langle \mathbf{x},\mathbf{x}\rangle}{\langle \mathbf{x},\mathbf{x}\rangle + 2\langle \mathbf{x},\mathbf{y}\rangle} + \frac{\langle \mathbf{y},\mathbf{y}\rangle}{\langle \mathbf{y},\mathbf{y}\rangle + 2\langle \mathbf{x},\mathbf{y}\rangle}\right).$$

The authors also provide a variation of the Edge score, which they call the HalfEdge score. Rather than considering edges, the HalfEdge score considers node-edge incidences, and asks the following: of all node-edge incidences involving a node of type x, what fraction of the relevant edges connect to a node also of type x? In the Edge score above, an edge between two x vertices would only be considered once; now, it's considered twice, because it is incident on two x vertices. The total number of edge-node incidences for an x-type vertex is given by $\langle \mathbf{x},\mathbf{x}\rangle + \langle \mathbf{x},\mathbf{y}\rangle$, and consequently the half-edge score is:

$$HalfEdge(\mathbf{x},\mathbf{y}) = \frac{1}{2}\left(\frac{\langle \mathbf{x},\mathbf{x}\rangle}{\langle \mathbf{x},\mathbf{x}\rangle + \langle \mathbf{x},\mathbf{y}\rangle} + \frac{\langle \mathbf{y},\mathbf{y}\rangle}{\langle \mathbf{y},\mathbf{y}\rangle + \langle \mathbf{x},\mathbf{y}\rangle}\right).$$

This has the intuitively appealing interpretation of the first term being the probability that a neighbor of an x person is another x person rather than a y person.

However, one drawback of the CAPY scores is that they ignore local adjacencies – By replacing each node with a complete graph, all people within a geographic unit are treated equally, as if they are all adjacent to each other. This means that individuals that live within a unit but far apart are treated the same as individuals that live next to each other. Areas that are disconnected in reality are now all related equally in the exploded graph, resulting in slight inaccuracies.

3.3.1 Example Outputs of CAPY

For this measure we will consider the Total Checkerboard Example (5c) with respect to the exploded dual graph. Consider the following equation:

$$\langle \mathbf{x},\mathbf{x} \rangle = \sum_i x_i^2 + \sum_{i \sim j} 2x_i x_j.$$

Intuitively, the first term is (double) counting the number of within-unit edges of the same group, and the second term is counting the number of between-unit edges of the same group. Due to the symmetry of this example, $\langle \mathbf{x},\mathbf{x} \rangle = \langle \mathbf{y},\mathbf{y} \rangle$. We can thus evaluate the following, noting there are six vertices each with x-population 10 and 3 edges that have x-population 10 at both endpoints,

$$\begin{aligned} \langle \mathbf{x},\mathbf{x} \rangle = \langle \mathbf{y},\mathbf{y} \rangle &= 6(10^2) + 3(2 \cdot 10 \cdot 10) \\ &= 1200. \end{aligned}$$

In consideration of $\langle \mathbf{x},\mathbf{y} \rangle$, the intuition is again that the first term is evaluating the number of within-unit edges of different groups, and the second term is the number of between-unit edges of different groups. As such the first term is always zero, as there are no within-unit edges between different groups in this example, and the second term reflects the fact that there are 17 between-node edges in the dual graph consisting of vertices of different groups. This example yields the following:

$$\langle \mathbf{x},\mathbf{y} \rangle = 0 + 17(10 \cdot 10) = 1700.$$

Thus, the CAPY Edge and Halfedge scores are as such:

$$\begin{aligned} Edge(\mathbf{x},\mathbf{y}) &= \frac{1}{2}\left(\frac{\langle \mathbf{x},\mathbf{x} \rangle}{\langle \mathbf{x},\mathbf{x} \rangle + 2\langle \mathbf{x},\mathbf{y} \rangle} + \frac{\langle \mathbf{y},\mathbf{y} \rangle}{\langle \mathbf{y},\mathbf{y} \rangle + 2\langle \mathbf{x},\mathbf{y} \rangle} \right) \\ &= \frac{1}{2}\left(\frac{1200}{1200 + 2(1700)} + \frac{1200}{1200 + 2(1700)} \right) \\ &= \frac{6}{23} \\ &= \boxed{0.2609}. \end{aligned}$$

$$\begin{aligned} HalfEdge(\mathbf{x},\mathbf{y}) &= \frac{1}{2}\left(\frac{\langle \mathbf{x},\mathbf{x} \rangle}{\langle \mathbf{x},\mathbf{x} \rangle + \langle \mathbf{x},\mathbf{y} \rangle} + \frac{\langle \mathbf{y},\mathbf{y} \rangle}{\langle \mathbf{y},\mathbf{y} \rangle + \langle \mathbf{x},\mathbf{y} \rangle} \right) \\ &= \frac{1}{2}\left(\frac{1200}{1200 + 1700} + \frac{1200}{1200 + 1700} \right) \\ &= \frac{12}{29} \\ &= \boxed{0.4138}. \end{aligned}$$

Example (5)	**Edge Output**	**HalfEdge Output**
Total Segregation	0.7059	0.8276
Gradual Segregation	0.3942	0.5655
Total Checkerboard	0.2609	0.4138
Partial Checkerboard	0.2264	0.3693
Random	0.3588	0.5217

Table 4: CAPY outputs for the examples introduced in Figure 5.

Outputs for CAPY scores for all examples are shown in Table 4.

Notice how we have no confirmation that the individuals residing in each unit are actually all connected – we may have inaccurate conclusions on adjacency as a result.

3.4 Random Walk Measure: Sousa and Nicosia (2022)

The last measure we will evaluate closely follows that of Sousa and Nicosia in their work on quantifying ethnic segregation in cities using random walks [19].

The method itself involves consideration of the dual graph for a region of interest. Each node contains information about some variable of interest – in the case of this problem, $\bar{x}_i$ is a vector containing the population of each racial group at node i. Rather than previous segregation measures which only consider two population subgroups, this method considers significantly more racial groups, including mixed-race groups, which is a more accurate reflection of our present-day world.

We are interested in the spatial distribution of $\bar{x}_i$, which will reveal to what extent nodes being spatially close in the graph also have similar values encoded in their vectors. By applying a random walk from any v_i (where each "step" goes to a neighboring node of v_i with equal probability), information is gathered as the walk goes on. The specific method proposed by Sousa and Nicosia measures the length of time t (in steps of the walk) it takes to reach all racial groups, which the authors refer to as classes. By "reaching a class" or "seeing a class," they mean visiting at least one node that has a non-zero population for that class. The idea is that greater values of t represent more segregated areas, and smaller values of t indicate more racial integration. The principal proposal of their paper is that the level of segregation of an area can be represented by Class Coverage Time (CCT) of a random walk on a corresponding dual graph G. CCT is defined as the expected number of steps needed by the random walker, starting at a random node, to visit some fraction c of total classes in the system.

Note that the previous measures all output a score between 0-1, with 1 being total segregation and 0 being total integration. In order to compare the random walk measure to the other scores, we will additionally compute the following metric: *If the random walk ends after all classes have been seen, what fraction of nodes have been visited?* Here a more integrated area would require a lower fraction of total nodes to be visited. A segregated region, or a region where all members of one particular class were clustered in a small corner of the graph, would require a much higher fraction of nodes to be visited. While this struck us as the most natural, intuitive way to convert Sousa and Nicosia's method into a single score that ranges between 0 and 1, it does suppress some information that the more general CCT calculations provide.

A drawback of this method is that it considers a racial group "seen" no matter how small it is, which could be only a few residents. This can be aided by choosing a higher threshold of class members needed to consider a class being seen, such as over 50%, etc. This method also has no strong basis for comparison to other segregation metrics, deeming it necessary to provide other metrics to contextualize what an output really means.

3.4.1 Example Outputs of Random Walks

Lastly, we will demonstrate how the Random Walk score was used to evaluate the Partial Checkerboard shown in example 5d. In this instance, no matter where it starts, the walk will always need only one step to see every class in the system: it sees both classes immediately. Therefore, we have a random walk output of $\frac{1}{12}$.

Exact values can easily be calculated in a similar way for the Gradual Segregation ($\frac{1}{12}$), Total Checkerboard ($\frac{1}{6}$), and Random $\left(\frac{10}{12} + 2 \cdot \frac{2}{12} = \frac{7}{72}\right)$ examples. For the Total Segregation examples, we ran 1000 trials and averaged the results. See Table 5 for the mean and standard error of the values produced in these trials, rounded to four decimal places.

Example (5)	**Output**	**Standard Error (1000 trials)**
Total Segregation (5a)	0.3789	0.0044
Gradual Segregation (5b)	$1/12 \approx 0.0833$	–
Total Checkerboard (5c)	$1/6 \approx 0.1667$	–
Partial Checkerboard (5d)	$1/12 \approx 0.0833$	–
Random (5e)	$7/72 \approx 0.0972$	–

Table 5: Random Walk outputs for the examples introduced in Figure 5.

Notice how this method has the same output for certain regions depending on how the classes are distributed, a lower bound of $\frac{1}{n}$ shown in the Gradual Segregation and Partial Checkerboard examples. This is most likely irrelevant when more classes are introduced into the system, as is the case in real-world applications. A potential downfall of this method is that outputs may differ across trials – this has little effect when the system is large and multiple trials are run, but shows up more noticeably in the small systems used in these examples.

3.5 Summary of Outputs and Basis for Chosen Method

In relation to previous mathematical measures of segregation, we believe that the explanatory power of the Random Walk method is somewhat higher than the other indices: It considers spatial relationships, changing aggregation levels, and within-unit adjacencies. The measure was chosen to evaluate the racial disparity of LA County (conducted in the next section) because it depends on the structural characteristic of the graph in question and the distribution of node properties, which therefore preserves the overall structure of the dual graph. This effectively allows researchers to compare the segregation of different systems on equal grounds, as the focus on preserving structure acts as a normalization aspect to relate contrasting levels of disparity across regions to one another. These aspects combine to let us analyze patterns in the structures of different regions, and objectively compare the overall levels of segregation across systems. Additionally, outputting specific lengths of time (rather than a single statistic) allows us to easily determine any glaring

outliers in the system. We also believe that the random walk measure contains the simplest method (fewest amount of calculations) to consider multiple racial groups instead of just two. As such this method complements our increasingly evolving society; it is extremely effective in adding and evaluating multiple racial groups.

However, we see many benefits to the other measures discussed, ultimately leading us to conclude that an area's racial segregation can be best understood by using multiple measures at once, and using each measure's strength to build a more comprehensive model. Table 6 displays the outputs for each score discussed, where considering the totality of the scores for a given example is more informative than any single score is on its own.

Example	**Moran's I**	**Dissimilarity**	**Edge**	**HalfEdge**	**Random Walk**
Total Segregation	0.7826	1	0.7059	0.8276	0.3
Gradual Segregation	0.7826	0.4	0.3942	0.5655	0.083
Total Checkerboard	0.2609	1	0.2609	0.4138	0.091667
Partial Checkerboard	0.2609	0.6	0.2264	0.3693	0.083
Random	0.5104	0.5749	0.3588	0.5217	0.1

Table 6: A table of outputs for each measure based on the examples introduced in Figure 5.

4 Applying Metrics to Los Angeles, CA

In this section we consider what the various segregation scores from the previous section tell us about Los Angeles. Rather than recomputing them, we note that in their paper on Clustering Propensity [2], Alvarez et al. provided Moran's I, Dissimilarity Index, and CAPY outputs for Los Angeles across three decades (whether it is Los Angeles County or the City of Los Angeles is not specified). We include them here in Table 7.

Year	**Moran's I**	**Dissimilarity**	**Edge**	**HalfEdge**
2010	0.546	0.523	0.473	0.633
2000	0.531	0.546	0.491	0.65
1990	0.604	0.55	0.497	0.663

Table 7: Outputs for various methods applied to Los Angeles over the span of three decades, as computed by Alvarez et al.[2].

We see that across all methods, the score variation across the three census years is at most ± 0.05, implying that there has been minimal change in the overall racial population distribution of Los Angeles in the past few decades. In the next subsections we provide our methods for computing the random walk score for 2020 data for Los Angeles County, and find it to be 0.6166 (due to computation and runtime limitations we were only able to consider data from one year, and chose to use the most recently available populations rather than older data).

The remainder of this section will focus on the methods, results, and interpretation of applying the Random Walk measure to LA County. In addition to the single random walk

score, we also provide a full CCT analysis and accompanying plots.

4.1 Data and Code

All data used in this analysis was originally collected from the U.S. Census database. Clean versions of the county-level data for the random walk analysis were provided by authors Sousa and Nicosia in their GitHub repository [17]. This data consisted of a list of class population information for each Census tract, and an edge-list describing which census tracts were geographically adjacent. Our results are built from a total of 3,923 Census tracts with 64 racial classes each.

The code used to analyze this problem was adapted from the authors [17]. Upon input (class population information for each census tract and an edge-list describing census tract adjacencies) the code records the total number of classes it has seen at each step and conducts a random walk until it has reached 100% of the total classes in the region. The resulting output is a vector of length 100: the first entry is the number of steps until the walker has seen 1% of the total classes in the system, the second entry the number of steps until the walker has seen 2% of the total classes in the system, ... , the nth entry is the number of steps until the walker has seen n% of the total classes in the system, until the final entry outputs the total number of steps needed to see every class in the system.

In what we refer to as one *trial*, this random walk process is done 3,923 times, once from each possible starting node. Each of these 3,923 random walks produced a length-100 vector, and we averaged these vectors entry-wise across all 3,923 walks. That is, the i^{th} entry in the averaged vector is the average number of classes seen after i steps across all possible starting nodes. For our results, we conducted 10 trials in order to provide a more accurate sense of scale for the possible spread of CCTs, due to the randomness in the random walks.

One key modification was implemented into the original code: updates were made such that the walk counts the classes seen in the starting node as being visited, rather than only recording after the first step; the code of Sousa and Nicosia did not count classes seen at the starting node. An earlier run[5] was calculated using the same method as Sousa and Nicosia, which produced slightly different results implying a higher level of racial segregation in LA County. We believe that the adaptation we make in this work is more representative of the goals of this method and therefore sustained this approach throughout our analysis (such that all results are internally consistent).

A complete documentation of the code and data used in this specific paper is available on GitHub[22].

4.2 Adapted Measure: Null Model

In order to have a basis from which to analyze the resulting output of Los Angeles County data, it was deemed necessary to implement a null model. Due to complications in recreating Sousa and Nicosia's null model, we ultimately created our own design that served the purposes needed in this analysis. We crafted a comprehensive null model as follows: original class-population data from each LA County Census tract was randomly permuted (across all 3,923 tracts) ten separate times. For each of these ten permutations, one trial (as described above) was performed, producing a length-100 vector of the average CCTs across all possible starting states. These vectors were then averaged entry-wise to give us the single, primary null model for CCTs used in our analysis.

By swapping the racial population counts randomly between nodes, we believe that this effectively eradicates any existing biases in the structure of LA County. This method of randomization provides us with a sufficiently arbitrary structure of comparison for LA County, and will help us determine how well Sousa and Nicosia's method effectively captures the geography of segregation in Los Angeles County.

4.3 Results of Random Walk Measure applied to Los Angeles County

4.3.1 Class Coverage Time Plots

Our output consists of a length-100 vector of average CCT for Los Angeles County, and a length-100 vector of average CCT for our null model, the n^{th} entry corresponding to the average number of classess seen after n steps. These are both shown in Figure 7, though the difference between them is difficult to see given the scale. We note that it takes around 3,580 more steps to reach 100% of the total classes in the null model vs. in LA County.

Because of the large range of values, and to assist in seeing the steepness of this curve, the natural log of the CCT is shown in Figure 8. The fact that this log plot is somewhat linear for most of its expanse suggests the original plot has approximately exponential growth. This also makes visible that it tends to take LA County a higher number of steps to reach the same percent of classes as the null model, for most of the range of percents considered. It is at around 80-90% of classes seen where their difference increases significantly.

We observe that the difference between the two models varies according to the percentage of classes seen, with a higher percentage corresponding to a larger difference and vice versa. We note that this is likely an effect of having the total fraction of classes being equal to 1, as the overall coverage time will increase when the system contains rare classes – as is the case in LA County, which has 64 different racial classes where multiple racial groups contain only a few citizens each.

4.3.2 Random Walk Score

Recall that we introduced a metric to more easily compare the random walk score to the other measures discussed, analyzing what fraction of the total number of nodes have been visited after seeing each class in the system. In this analysis, data from the LA County walk visited 61.7% of all nodes in order to encounter all racial classes, and the null model visited 61.3% on average across sub-null models.

We can see that as these percentages are extremely similar, it suggests that the random walk score does not effectively capture the spatiality of segregation: randomly permuting the populations of nodes (and their racial compositions) has an extremely small effect on this random walk score. Additionally, though it may be useful in contextualizing different scores with respect to one another, we assert that attempting to summarize segregation in a single score can be ineffective when wanting a more comprehensive picture of racial segregation.

4.3.3 Spread of Class Coverage Time

Because the null model is comprised of the average of 10 sub-null models, we consider the spread across these 10 models and compare it to the values seen in LA County to more deeply assess the differences between the null model and LA county. For the sake of concreteness we focus on the class coverage times near 50% of classes seen, but saw a

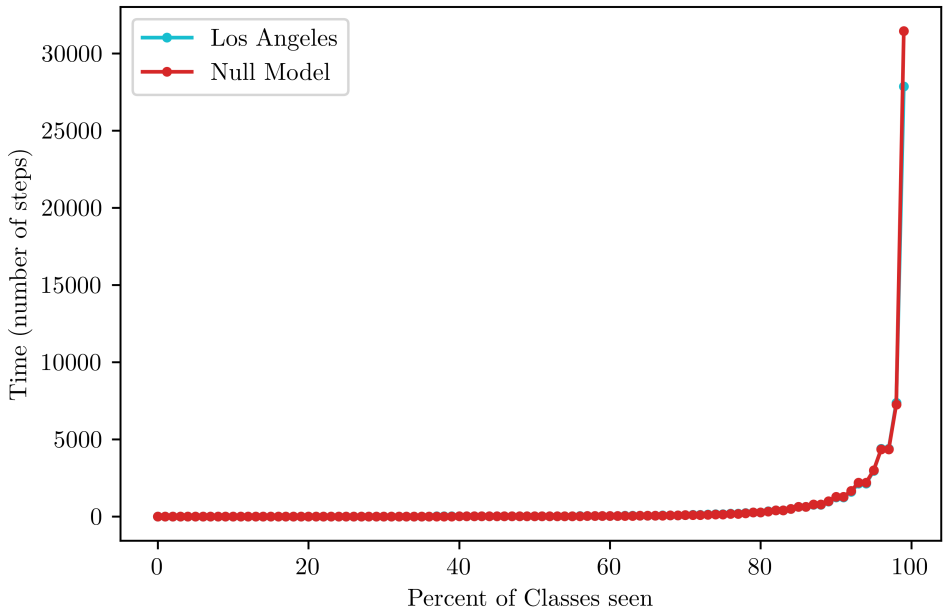


Figure 7: Complete output of CCT for LA County and the null model

similar picture throughout the range where LA county had longer coverage times than the null model. See Figure 9, which visualizes the spread of the class coverage times across the 10 sub-null models considered.

The minimum value seen across all ten sub-null models is 10.6 steps for 50% and 12.4 steps for 51%, while the maximum is 10.8 steps for 50% and 12.7 steps for 51%. With respect to the averages, the primary null model is 10.8 steps at 50% and 12.6 steps at 51%, compared to LA County's 14.2 steps and 16.4 steps, respectively. Notice that the deviation between sub-null models and the primary null model here is at most $\pm 0.2 \implies \pm 2\%$ (compared to the 34-37% deviation between LA County and the null model). This indicates that there are minimal levels of variation across early sub-null model runs, which also implies that LA is consistently at higher values than all the null models across the range of trials, and not solely the average.

Notice how Figure 9 exhibits step function-like growth. This is because, despite there being 64 racial classes, this work considers integer percentages: For instance when 31 out of 64 classes have been seen this is 48.4% of classes, and when 32 out of 64 classes have been seen, this is 50% of classes. It is in the same moment that 49% of classes have been seen that 50% of classes have also been seen when defining percentages this way. However, this step-function like growth is only visible in extremely zoomed-in figures, and has minimal affects on the overall high-level trends seen.

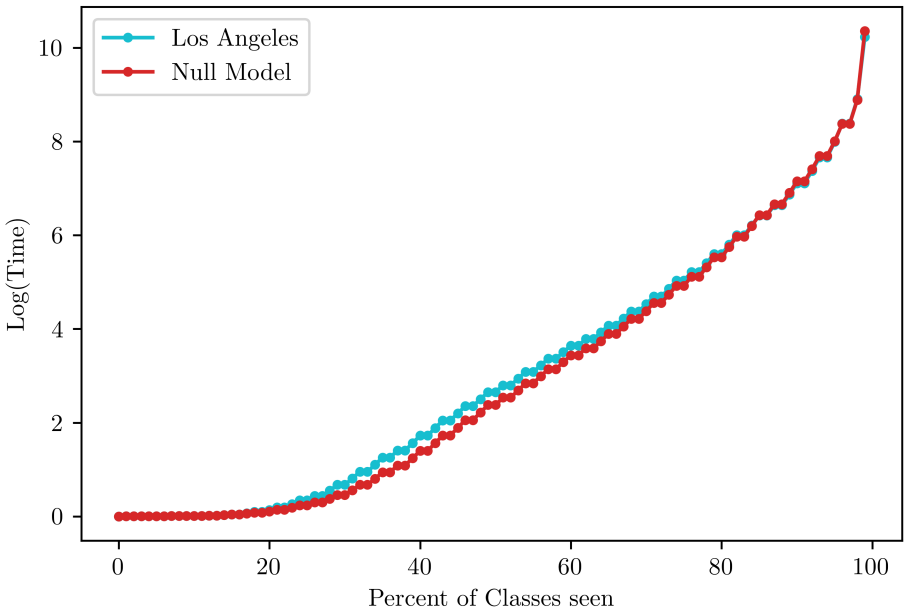


Figure 8: Natural log of the CCT for LA County and the null model

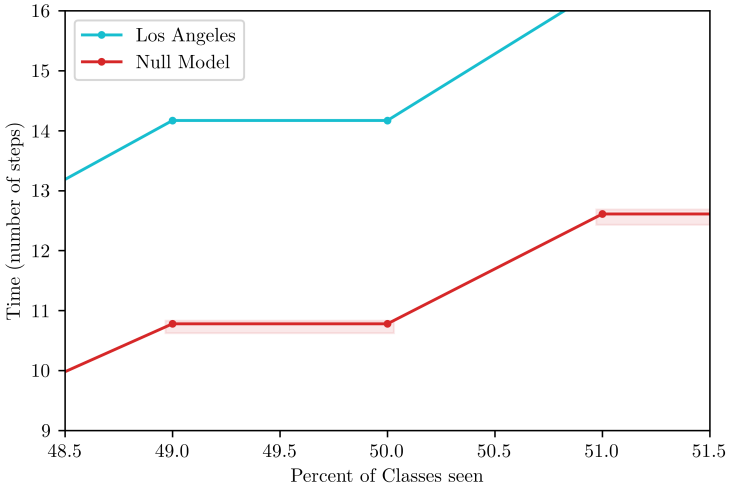


Figure 9: CCT for LA County and the null model, where the shaded bands indicate the min/max ranges across the sub-null model runs. Note that the range of sub-null models has only a slight deviation from the primary null model, which is their average.

4.3.4 Spatial Heterogeneity

In their analysis, Sousa and Nicosia introduced the concept of spatial heterogeneity ($\Delta\mu$): The notation $\mu(c)$ is used to denote the average number of steps required to see a c fraction of classes. $\Delta\mu$ is then the average deviation between a given region ($\mu(c)$) and its corresponding null model ($\mu^{null}(c)$), as follows:

$$\Delta\mu = \int_0^1 |\mu(c) - \mu^{null}(c)|dc.$$

That is, $\Delta\mu$ integrates to find the total area between the CCT curves for a given region and its null model.

Though we constructed a different null model than in their analysis, we applied the same approach to quantify the difference between our primary null model and LA County. Given the nature of the equation, both regions when $\mu(c)$ is larger than $\mu(c)^{null}$ and regions when the opposite is true contribute equally to this score. Approaching this measurement without the use of absolute values is a possible subject for future analysis.

We computed the average deviation as a discrete summation over the length-100 output vectors describing our CCT:

$$\Delta\mu = \sum_{k=1}^{100} \frac{1}{100} \left| \mu\left(\frac{k}{100}\right) - \mu^{null}\left(\frac{k}{100}\right) \right|.$$

For the deviation between LA County and its null model, we calculated a spatial heterogeneity score of $\Delta\mu = 45.8971$. However, because our null model was created differently, we are unable to compare this value to those in Sousa and Nicosia's analysis in any meaningful way. Instead, we hope to supply additional context by providing spatial heterogeneity scores between the 10 sub-null models that went into the eventual primary null model.

In order to analyze each difference pairwise, we calculated across the $\binom{10}{2} = 45$ different scores. For $1 \leq i < j \leq 10$, let $\mu_i(c)$ and $\mu_j(c)$ be the average number of steps required to see a c fraction of classes in sub-null models i and j, respectively. We calculated the spatial heterogeneity scores for each pair of sub-null models as above:

$$\Delta\mu_{ij} = \sum_{k=1}^{100} \frac{1}{100} \left| \mu_i\left(\frac{k}{100}\right) - \mu_j\left(\frac{k}{100}\right) \right|.$$

The average across these scores was then calculated:

$$\Delta\mu_{null} = \frac{1}{45} \sum_{1 \leq i < j \leq 10} \Delta\mu_{ij}$$

resulting in a spatial heterogeneity value of $\Delta\mu_{null} = 80.92$, which is almost twice the size of LA County's spatial heterogeneity score, suggesting that LA County falls within the range of possibilities spanned by the sub-null models.

However, we suspect the main contributors to this score come from the large values of c close to 1, and thus this single score is not the most representative way to describe the relationships between different models. Visualization of differences is shown in Figure 10 as the natural log of the difference values ($\mu(c) - \mu^{null}(c)$) along with the natural log of the average difference across all sub-null models. The visualization shows that the difference

between LA County and the null model tends to be larger than the average difference between each of the sub-null models. The results of this section indicate a higher level of racial segregation in LA County than in the null model at most instances of the walk, suggesting the model meaningfully captures an important feature of LA county for this range.

However, the average difference between sub-null models is higher than the difference between LA County and the null model between 82-100% of total classes seen; when these differences are averaged not on a log scale (as the spatial heterogeneity score does), they make a much larger contribution to the score and are thus likely the reason the spatial heterogeneity score between the null models is so much higher than between LA County and the null model. The high spatial heterogeneity score between sub-null models could also possibly be due to the adaptation of starting states mentioned previously: our sub-null models now contain a higher spread given that we removed an element of randomness to the procedure. Previous outputs (see [5]) had a tighter concentration around the mean, resulting in a smaller spatial heterogeneity score across sub-null models.

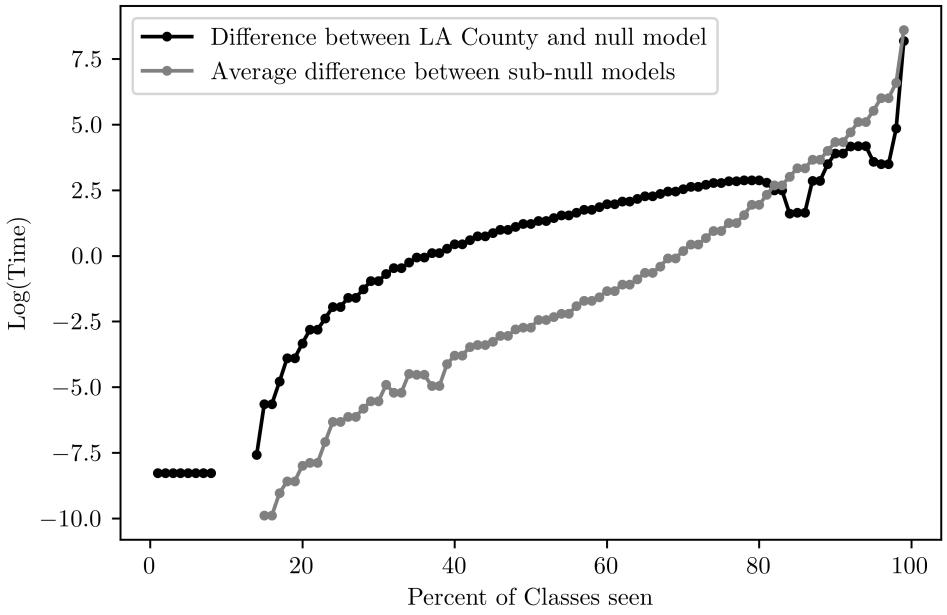


Figure 10: Natural log of the difference between LA County and the null model and the average difference between sub-null models.

5 Conclusion

The work and results from this study provide a sufficient basis for concluding that using a single score to measure segregation is insufficient, and that using multiple assessments is preferred to more accurately evaluate this issue and avoid erasing critical information. More research is needed to create a sufficiently representative way of quantitatively contextualizing racial segregation – This includes the development of additional scores that can complement those presented here in order to give a more complete picture.

In evaluating the different measures, we found many benefits and drawbacks to each: Moran's I accounts for spatial relationships between units, but changes based on how the units are defined. It focuses on adjacencies between units and ignores those within units – given the age of this score, problems with it are well-known across researchers in the field. The Dissimilarity Index does not account for these crucial spatial relationships, but it does offer a strong basis of comparison. This score is relatively simple compared to the others, which is helpful in terms of accessibility of information. The Clustering Propensity score is much less influenced by how units are defined and considers adjacencies within units, but ignores/fabricates local adjacencies. Assuming each resident in a unit is adjacent to every other resident in the unit is unrealistic. Lastly, the Random Walk score relies heavily on local adjacencies, but does not easily standardize to the [0,1] range as do the other scores. It is extremely efficient in evaluating multiple racial groups, but only considers whether a racial group is present in each unit rather than size of the group within that unit. It also takes more work to create a proper basis of comparison.

The work done here is just a fraction of the possibilities available towards creating a more just and equitable society, and readers are encouraged to continue their knowledge of intersectional issues spanning race, political structures, and quantitative analysis. Being able to objectively provide evidence that there is inequality present in a system in a crucial step towards eradicating it, and this cannot be done without acute levels of awareness across fields and a desire to enact change.

Acknowledgements. S. Cannon is supported in part by NSF grant CCF-2104795. An earlier version of this paper was Z. Dhillon's Undergraduate Senior Thesis at Claremont McKenna College; this version is available at `https://scholarship.claremont.edu/cmc_theses/3341/`. The authors can be reached at: `scannon@cmc.edu`, `zdhillon23@cmc.edu`.

References

[1] Aboelata, Manal J. Healing LA Neighborhoods: A once-in-a-generation opportunity to create thriving and inclusive communities across Los Angeles. *Preventioninstitute.org*, 2020. Available at `https://www.preventioninstitute.org/publications/healing-la-neighborhoods-once-generation-opportunity-create-thriving-and-inclusive`.

[2] Alvarez, Emilia, Duchin, Moon, Meike, Everett, and Mueller, Marshall. Clustering propensity: A mathematical framework for measuring segregation. *preprint*, 2018. Available at `https://mggg.org/Capy.pdf`.

[3] Ballotpedia. Large counties in the United States by population. `https://ballotpedia.org/Large_counties_in_the_United_States_by_population`, 2021.

[4] Charles, Camille Zubrinksy. *Won't you be my neighbor: Race, class, and residence in Los Angeles*. Russell Sage Foundation, 2006.

[5] Dhillon, Zarina Kismet. Measuring Racial Segregation in Los Angeles County using Random Walks, 2023. Available at https://scholarship.claremont.edu/cmc_theses/3341.

[6] Moon Duchin and Olivia Walch. *Political Geometry*. Springer, 2021. Available at https://mggg.org/gerrybook.html.

[7] Duncan, Otis Dudley and Duncan, Beverly. A methodological analysis of segregation indexes. *American sociological review*, 20(2):210–217, 1955.

[8] Equiano, Olaudah. *The interesting narrative of the life of Olaudah Equiano, or Gustavus Vassa, the Africian.* Norwich, 1789. Retrieved from the Library of Congress, https://www.loc.gov/item/44015764/..

[9] Federal Housing Administration. Underwriting Manual: Underwriting Analysis under Title II, Section 203 of the National Housing Act, 1936.

[10] Franklin, John Hope. History of Racial Segregation in the United States. *The ANNALS of the American Academy of Political and Social Science*, 304(1):1–9, Mar 1956.

[11] Lincoln, Abraham. 13th Amendment to the U.S. Constitution: Abolition of Slavery. *National Archives*, 1865. Available at https://constitution.congress.gov/constitution/amendment-13/.

[12] Mapping Inequality. Redlining in New Deal America. https://dsl.richmond.edu/panorama/redlining/#loc=10/34.005/-118.486&maps=0&city=los-angeles-ca, Accessed 2023.

[13] Moran, Patrick AP. Notes on Continuous Stochastic Phenomena. *Biometrika*, 37(1/2):17–23, 1950.

[14] National Association of Real Estate Boards. "Code of Ethics". https://www.nar.realtor/about-nar/history/1924-code-of-ethics, 1924.

[15] Redford, Laura. The Intertwined History of Class and Race Segregation in Los Angeles. *Journal of Planning History*, 16(4):305–322, 2017.

[16] Rodriguez-Moral, Antonio and Vorsatz, Marc. An overview of the measurement of segregation: classical approaches and social network analysis. *Complex networks and dynamics: Social and economic interactions*, pages 93–119, 2016.

[17] @sandrofsousa. GitHub - segregation-rw/ethnic-segregation-rw: Software to compute segregation based on random walks as in the paper "Quantifying ethnic segregation in cities through random walks". https://github.com/segregation-rw/ethnic-segregation-rw, Jul 2022.

[18] Smyton, Robin. How Racial Segregation and Policing Intersect in America. *Tufts Now*, 17, 2020. Available at https://now.tufts.edu/2020/06/17/how-racial-segregation-and-policing-intersect-america#:~:text=In%20short%2C%20the%20police%20reproduced,in%20racial%20exclusion%20and%20control.

[19] Sousa, Sandro and Nicosia, Vincenzo. Quantifying ethnic segregation in cities through random walks. *Nature Communications*, 13(1):5809, 2022. Available at https://doi.org/10.1038/s41467-022-33344-3.

[20] Taeuber, Karl E and Taeuber, Alma F. *Residential segregation and neighborhood change*. Transaction Publishers, 2008.

[21] U.S. Census. Los Angeles County, California - U.S. Census Bureau QuickFacts. https://www.census.gov/quickfacts/fact/table/losangelescountycalifornia/RHI125221#RHI125221, 2022.

[22] @zarinad. GitHub - zarinad/Random-Walks-in-LA-County: Code and data used in Sarah Cannon and Zarina Dhillon's "Evaluating Methods used to Quanitfy Racial Segregation". https://github.com/zarinad/Random-Walks-in-LA-County, July 2023.