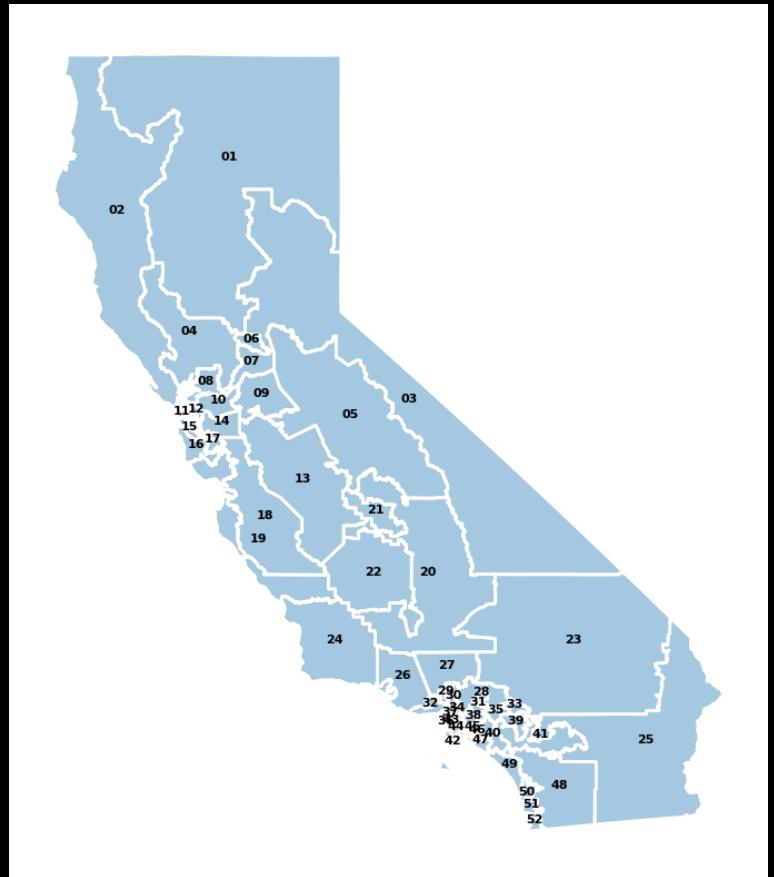


ANGELA ZHOU

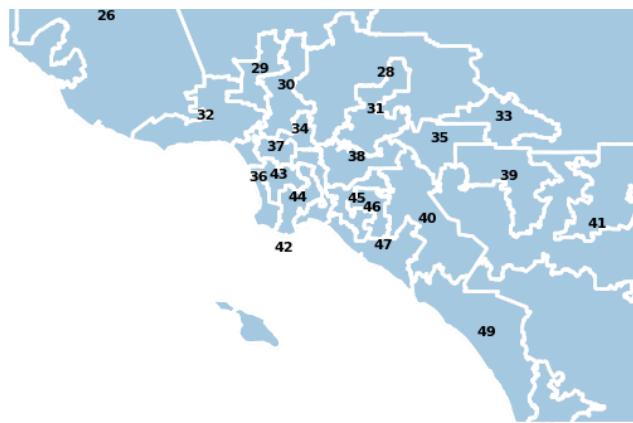
# MATH OF POLITICAL DISTRICTING

MATH195CM FALL 24 FINAL PROJECT  
WITH PROFESSOR SARAH CANNON



# DO YOU KNOW YOUR DISTRICTS?

California's 28th  
Congressional District

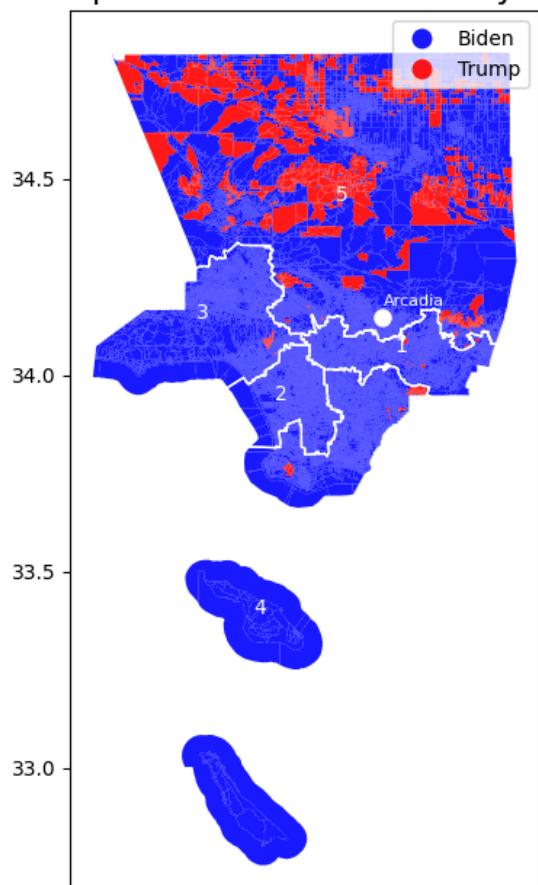


**Do you think the way these  
are divided make sense?**

01

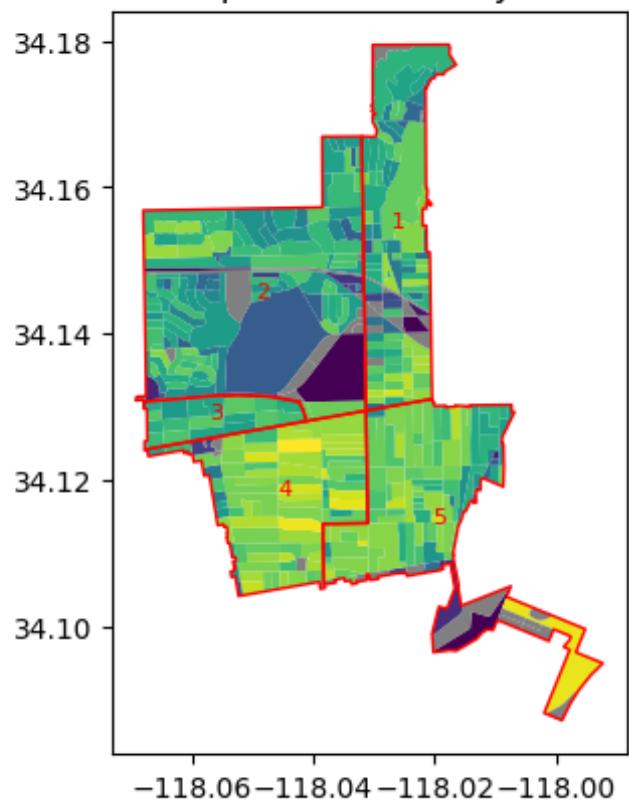
LA County Districts

LA County Trump vs. Biden Dominance by District in 2020



Arcadia City Council Districts

Arcadia Asian Population with City Council Districts



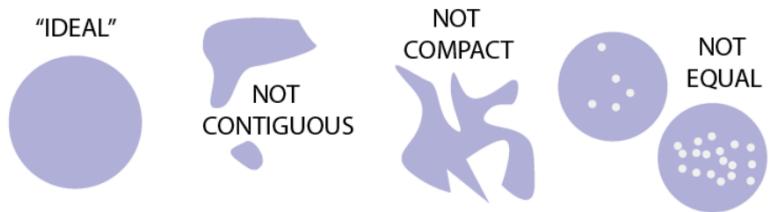
# HOW ARE DISTRICTS DRAWN?

In a process called **REDISTRICTING**, states, counties, and cities redraw boundaries for voting areas based on different factors like population and representation. Each state has their own process for redrawing.

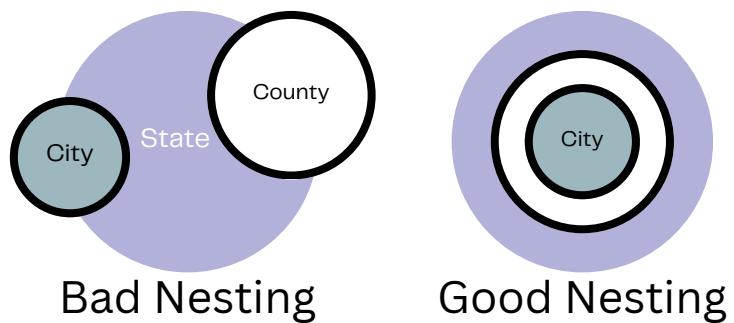
## CALIFORNIA

- California is divided into 52 Congressional districts that each elects a House representative
- Every 10 years, after the federal government updates the population count (**CENSUS**), California redraws its districts
- California has an **independent commission** to draw districts, created by voters in 2008 with members selected from the community. They include 5 Republicans, 5 democrats, and 4 with no party affiliation.
- New districts has to follow **rules** —————
- The commission considers public input. You can get involved by drawing your own map, submitting a comment, calling in, or joining advocacy groups.

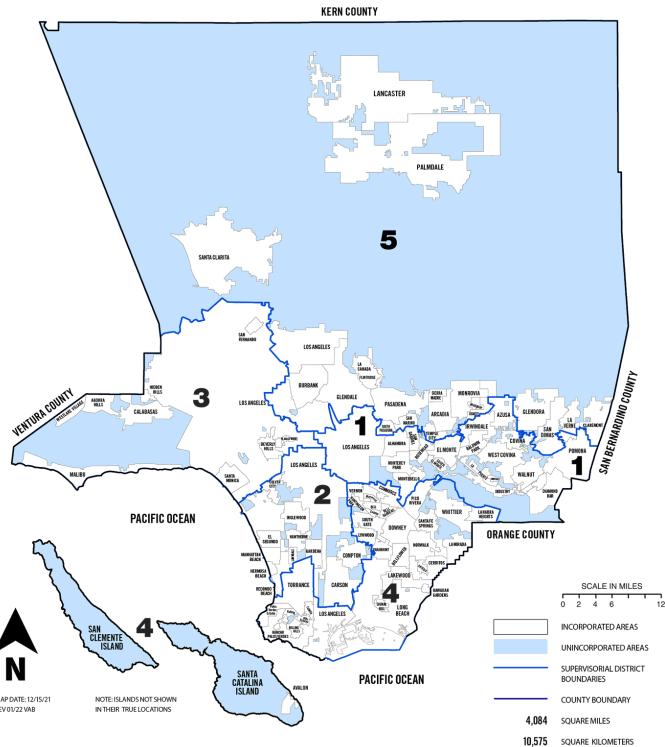
- 1. Equal population between districts
- 2. **Voting Rights Act**: Minority groups have to have equal opportunity in elections
- 3. **Contiguity**: Districts have to be connected



- 4. Keep together **communities of interest**: groups of people who live near each other and share common social or economic interests (culture, jobs, language)
- 5. **Compactness**: The shape of the district should be “round”
- 6. **Nesting**: Districts at a local level (county, city) should be inside the bigger statewide level district

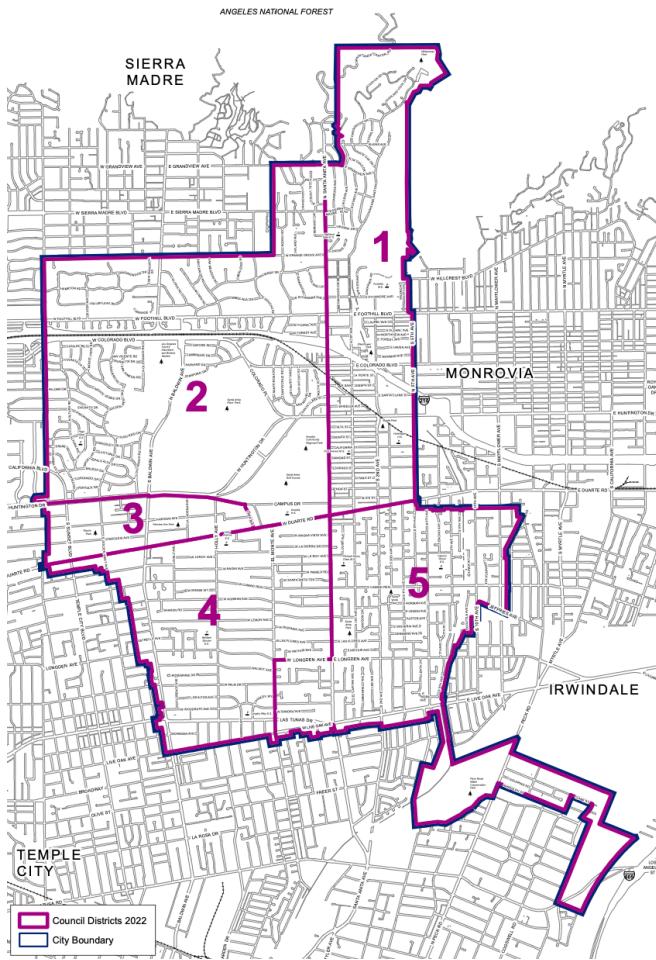


# HOW ARE DISTRICTS DRAWN?



## LA COUNTY

- This plan was drawn by the independent Citizens Redistricting Commission' (all LA County residents) in 2021 following the same rules above
- With 5 districts, it represents LA County's 88 cities and communities in LA County Supervisor elections
- The commission's goal was to give Latino and Asian residents better representation while maintaining a concentration of Black voters



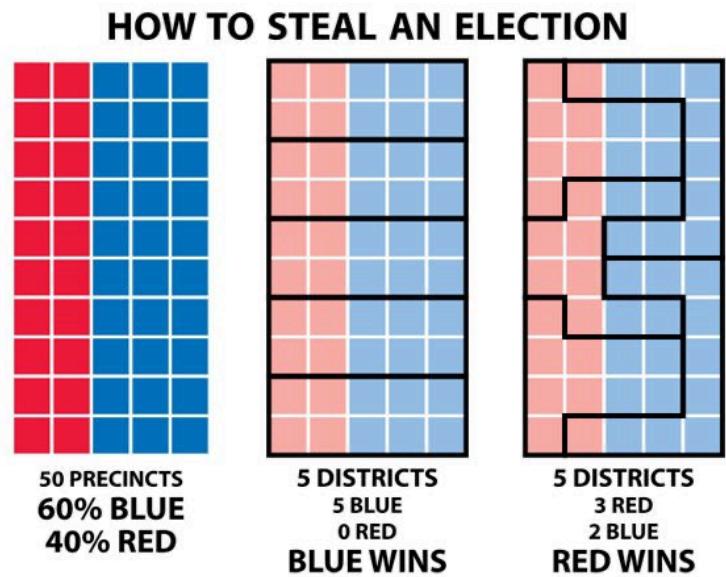
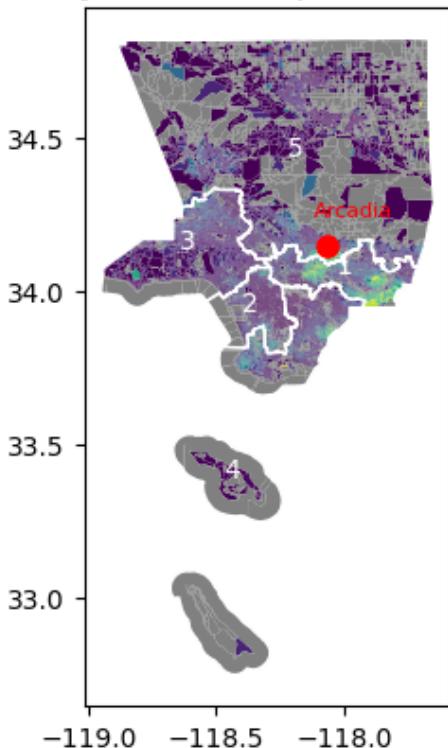
## ARCADIA

- The city of Arcadia is divided into 5 districts, each with an elected City Council Member
- With public input, the City Council redraws districts every 10 years following the same rules as above and while following natural boundaries like rivers, streets, etc. and without favoring or discriminating against a political party

# WHAT IS GERRYMANDERING?

Gerrymandering is when district boundaries are purposely drawn for a specific political result, like packing voters together to fewer districts or spreading voters out to multiple districts to limit voter power

LA County Asian Population Fraction



## WHY SHOULD YOU CARE?

The way districts are drawn should represent communities. For example, LA County is 14.7% Asian while Arcadia is 57% Asian. Do you feel like this is represented in your government leaders? Districts give communities voting power through representation and issues that affect daily life from a local level to big elections every 4 years.

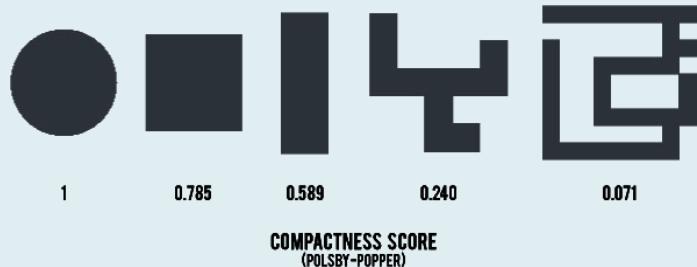
# HOW DO NUMBERS HELP IDENTIFY GERRYMANDERING?

## COMPACTNESS

There are a lot of scores that measure compactness by looking at a region's shape. For example, the **Polsby-Popper score**

**Polsby-Popper score** compares regions to a circle so that rounder districts are better than thin, long districts.

However, only looking at the shape can create problems when physical geography like rivers are good boundary lines but lead to bad scores. Many scores also don't consider people or population.



## PARTY FAIRNESS

There are also ways to detect **partisan gerrymandering**, where one party has an advantage because of districts. But scores often aren't enough to be a complete measure of fairness because they assume voter behavior to make predictions.

**Proportionality:** The number of seats a party wins should match the percentage of votes received. 50% of votes = 50% of seats. However, in a competitive district, a small change towards preferring a party can lead to a big change in proportionality.

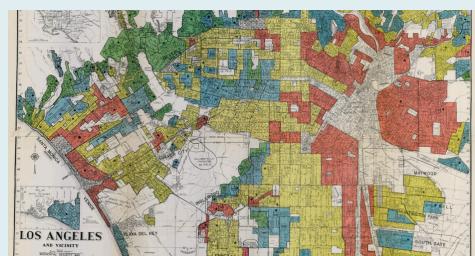
**Partisan Symmetry:** If parties got the same percentage of votes, they should get the same number of seats. If Democrats got 60% of the vote and received 5/10 seats, then in another election if Republicans got 60% of the vote, they should also receive 5/10 seats.

**Low Efficiency Gap:** Parties should be “wasting” the same number of votes (votes not needed to win or lose an election.)

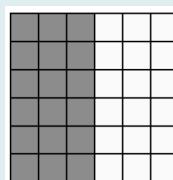
## SEGREGATION

Historically, the United States has created segregation through **redlining**, an unfair way of using mortgage, home loans, and more to restrict immigrant and Black populations to specific areas and lower incomes. In LA, redlining has consequences like the homelessness crisis from lack of housing, especially in areas like Highland Park and Boyle Heights.

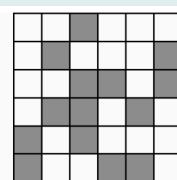
**Moran's I** is a score to measure segregation by checking how separated, evenly mixed, or clustered a district's population is.



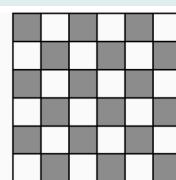
A 1939 map of segregation and redlining in LA



Positive, segregation is occurring



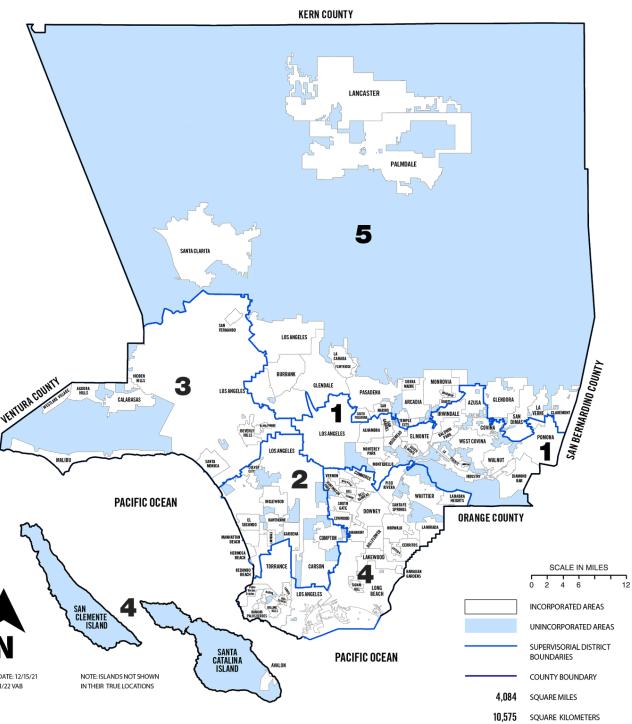
Randomly dispersed



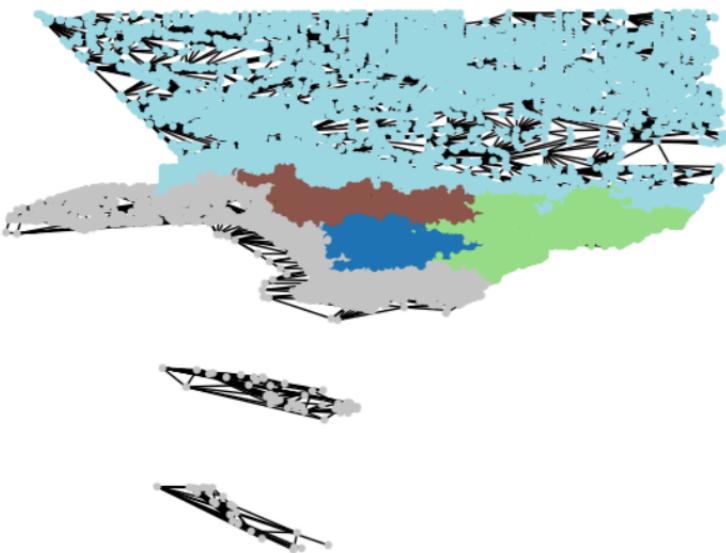
Negative - intentionally distributed

# WHAT ABOUT USING COMPUTERS?

In districting, maps are turned into **graphs** to represent how voting areas are connected and help plan how districts can be grouped. Graphs are made of dots and connecting lines, where dots are towns, or districts, and lines show which districts are neighbors and next to each other.



LA County map turned into a randomly sampled graph with different districts



Because there's so many different possible plans, there needs to be ways of creating and comparing lots of different plans. **Random sampling** is a method of randomly picking lots of possible maps and seeing how fair or balanced they are. **Random walks** is a way of wandering step by step in random directions on a graph, and used in districting to explore different ways of drawing districts through making small changes. A **Markov chain** uses these random walks to go from one plan to another based on probabilities and set rules, like population count or compactness.

# WHAT ABOUT USING COMPUTERS?

**ReCom**, or recombination, uses random walks and markov chains to make new districts specifically by combining two districts and randomly splitting them to make a new plan. at every **step** in the walk.

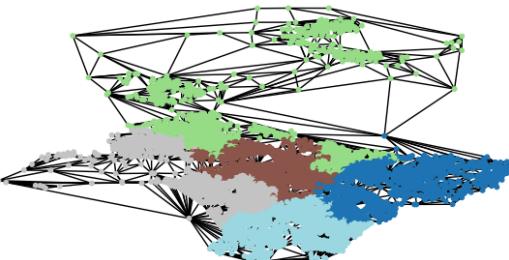
**Ensemble Analysis** is a general process of building a group of plans and comparing it to current plans.

**Gerrychain** is the specific tool in code that includes ReCom, which I've used along with ensemble analysis to analyze LA County's districts.

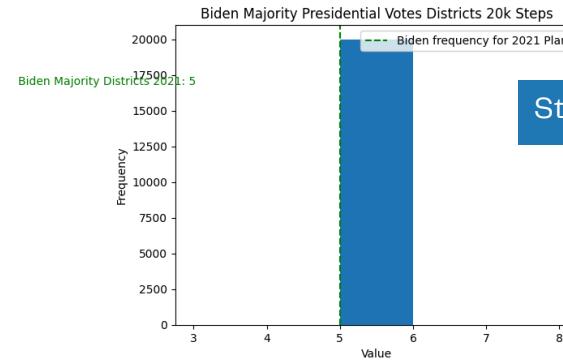
1. I turned a map of LA County that had data about population, race, voting history, and the county districts into a **graph**.
2. I made a **random** districting plan
3. I ran **ReCom** from my random plan for 10,000 and 20,000 **steps**.
4. I used **ensemble analysis** to compare the average Asian population and majority presidential votes for Biden in my generated plans with 2021's plan.
5. I also compared 10,000 steps to 20,000 to each other to make sure I had ran it for enough steps

I found that the generated plans and 2021's plan both had a majority of all 5 districts vote for Biden. They had a similar count of the Asian population, while none of the districts had an Asian majority. **Since the generated plans and 2021's plan are similar, there isn't extreme partisan or racial gerrymandering.**

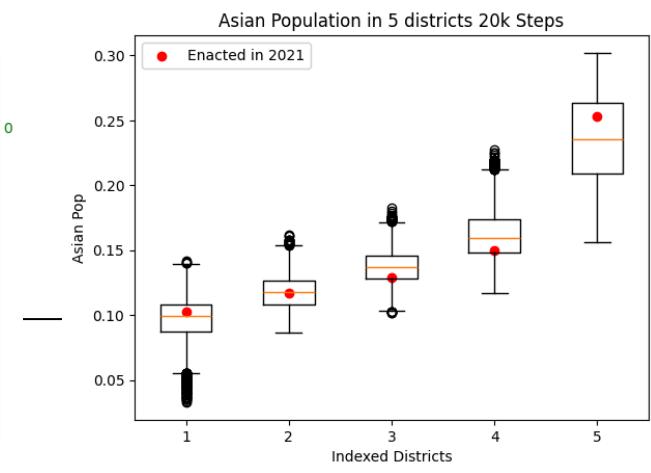
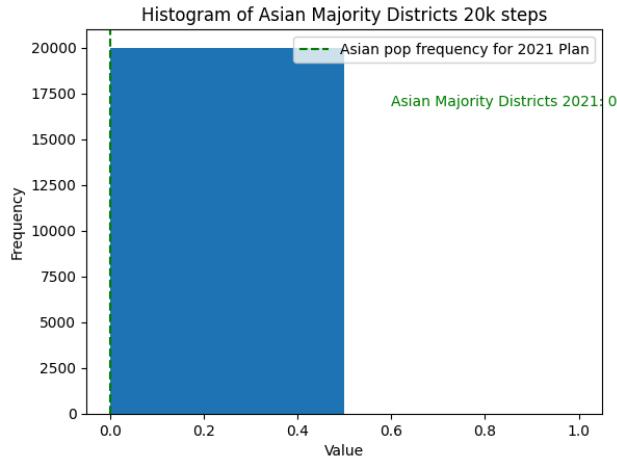
LA County Initial Districting Plan



Step 1 and 2



Step 4, ensemble analysis



# CHALLENGES

There are a lot of challenges when it comes to using math or computers to try to solve redistricting problems.

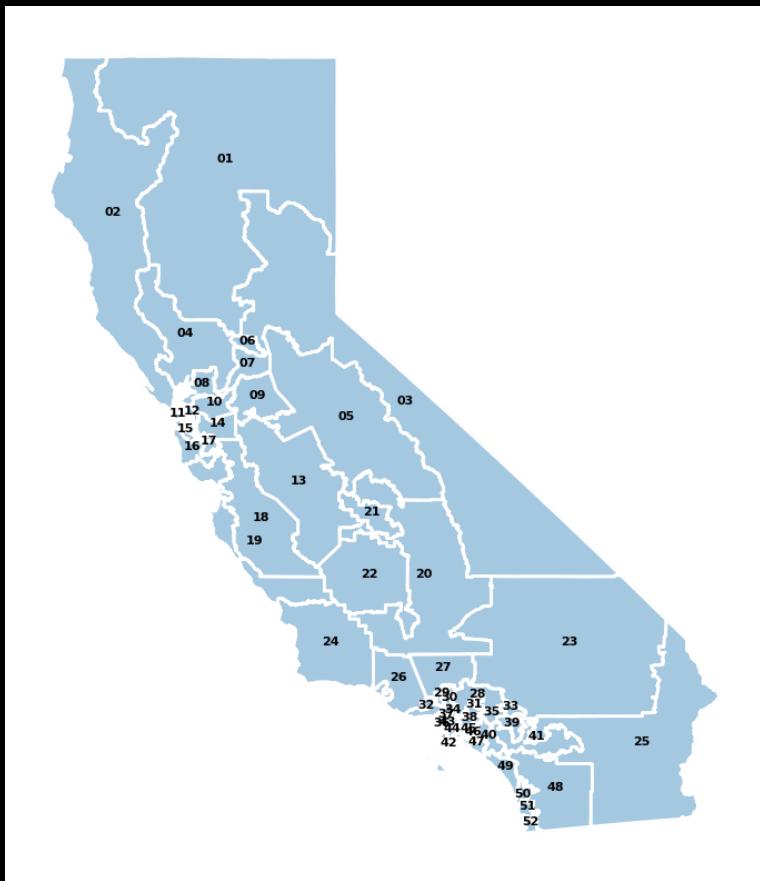
1. Getting accurate, standardized data is hard especially when it's data about people or geography. Data can be sensitive to politics, boundaries, physical geography like water, and more.
2. Numbers and computers are a good place to start when measuring fairness or looking for gerrymandering, but should not be the only solution. Numbers and computers can also be biased and limited as they don't take always into account important histories or people - which is always changing from place to place.

# THANK YOU!

azhou4013@scrippscollege.edu

Code:

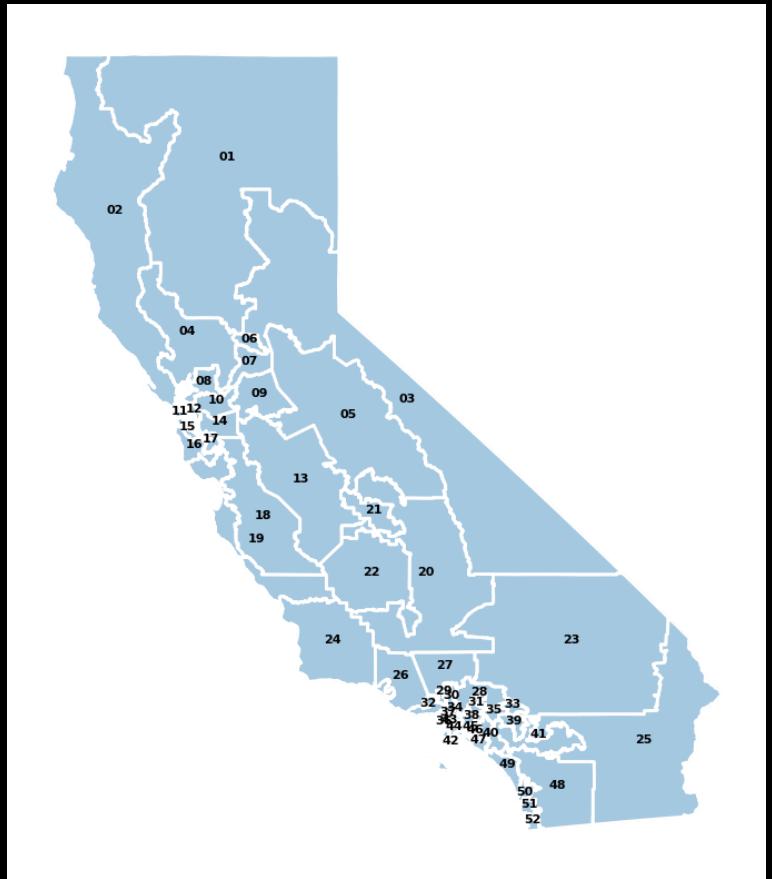
Technical Appendix and sources:



ANGELA ZHOU

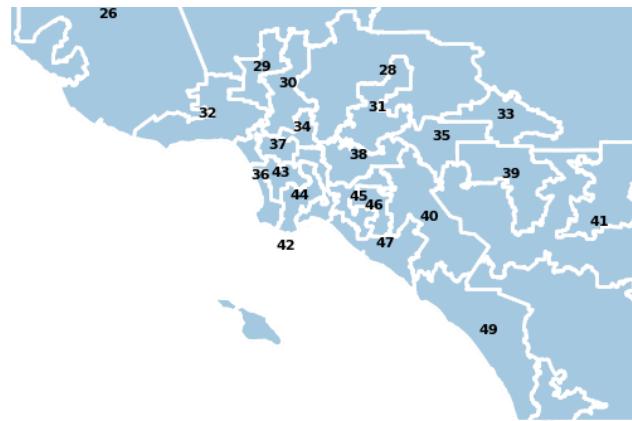
# 政治分区划分的数学

MATH195CM FALL 24 最终项目和教授  
SARAH CANNON



# 您了解您的地区吗？

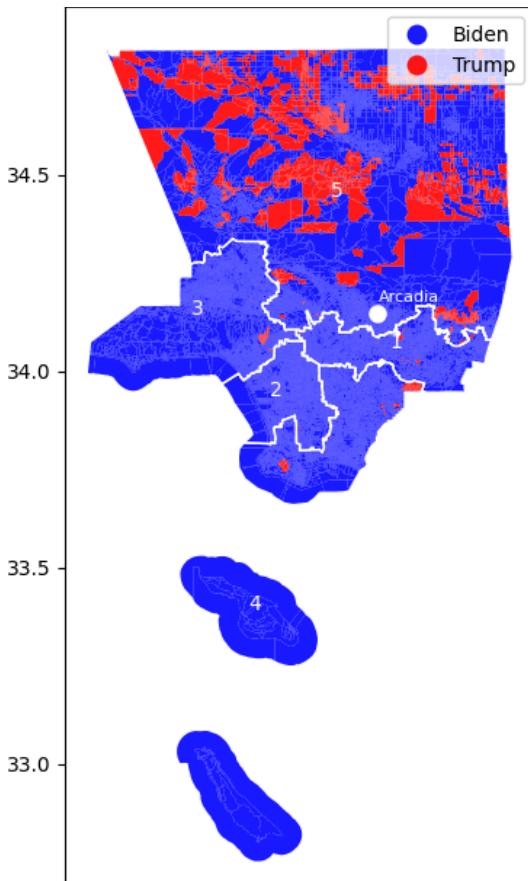
## 加州第 28 国会选区



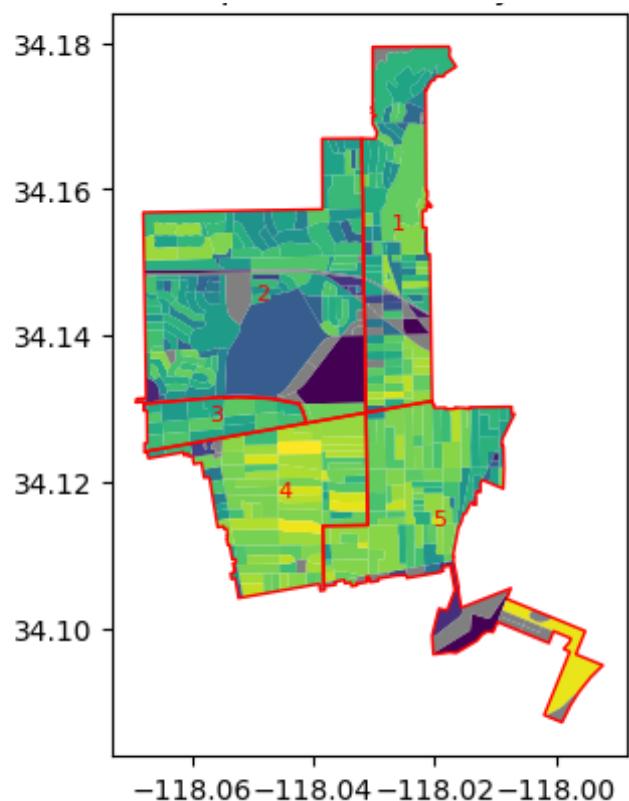
您认为这样的划分方式合理吗？

01

洛杉矶县选区，特朗普与拜登在 2020 年的主导地位



亚裔人口占多数的 Arcadia 市议会选区



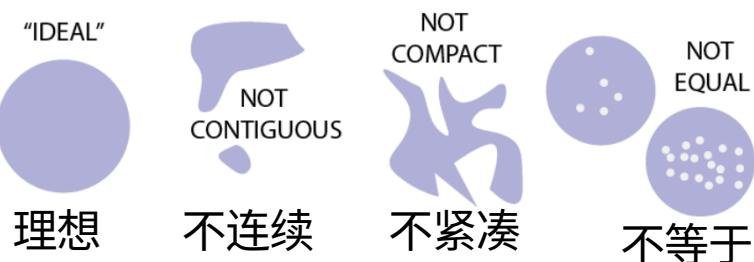
# 区域是如何划分的？

在重新划分选区的过程中，各州、县和市会根据人口和代表人数等不同因素重新划分选区边界。每个州都有自己的重新划分选区程序。

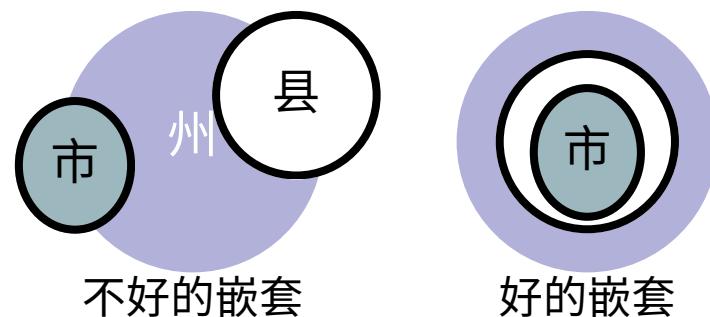
## 加州

- 加州分为 52 个国会选区，每个选区选举一名众议院代表
- 每 10 年，在联邦政府更新人口数量(**CENSUS**)后，加州都会重新划分选区
- 加州有一个独立委员会负责划分选区，该委员会由选民于 2008 年选出，成员均来自社区。委员会成员包括 5 名共和党人、5 名民主党人和 4 名无党派人士。
- 新区必须遵守规则——
- 该委员会会考虑公众的意见。您可以通过绘制自己的地图、提交评论、致电或加入倡导团体来参与。

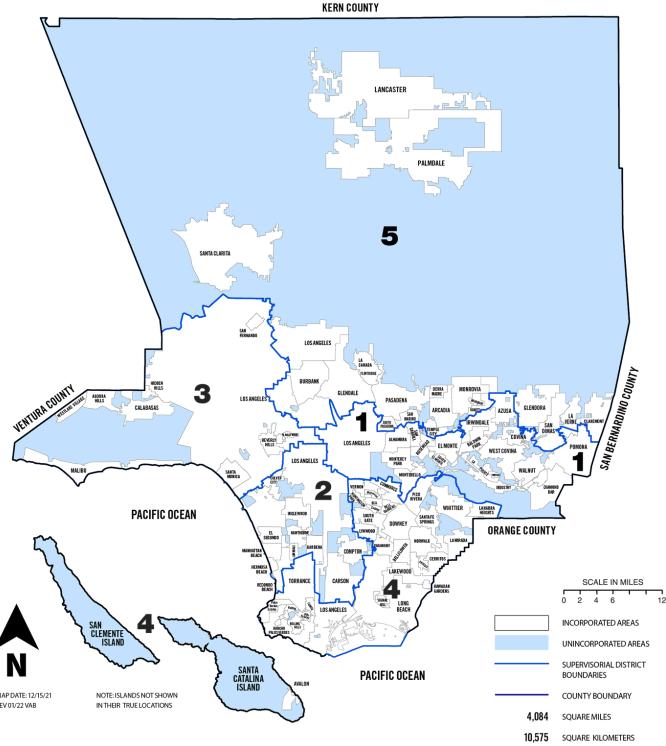
1. 各地区人口均等
2. **投票权法案, Voting Rights Act**: 少数群体在选举中必须享有平等机会
3. **毗连性, Contiguity**: 各区域必须连通
4. 保持**利益共同体**: 居住在一起并拥有共同社会或经济利益（文化、工作、语言）的人群
5. **紧凑, Compactness**: 区形宜“圆”
6. **嵌套, Nesting**: 地方级别（县、市）的区应该位于更大的州级区内



理想 不连续 不紧凑 不等于

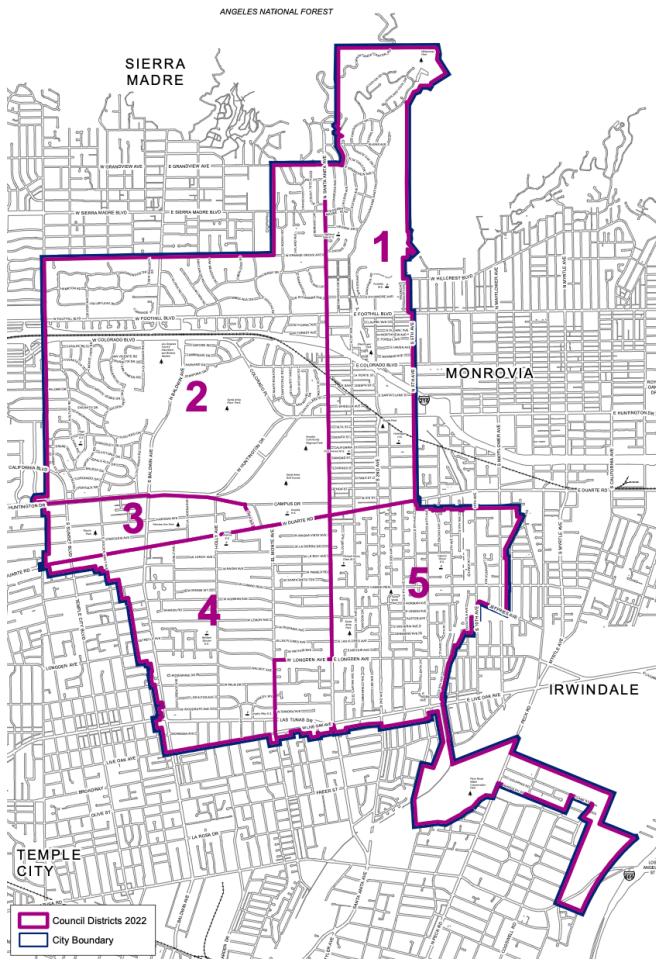


# 区域是如何划分的？



## 洛杉矶县

- 该计划由独立公民选区重划委员会（所有洛杉矶县居民）于 2021 年制定，遵循上述相同规则
- 它拥有 5 个选区，代表洛杉矶县的 88 个城市和社区参加洛杉矶县监事选举
- 该委员会的目标是让拉丁裔和亚裔居民获得更好的代表权，同时保持黑人选民的集中度



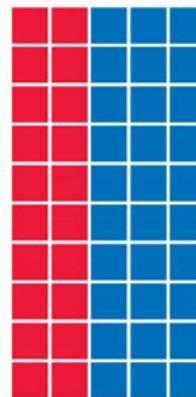
## ARCADIA

- 阿卡迪亚市分为 5 个区，每个区都有一名民选市议会议员
- 市议会根据公众意见，每 10 年重新分区，遵循上述相同规则，同时遵循河流、街道等自然边界，不偏袒或歧视任何政党

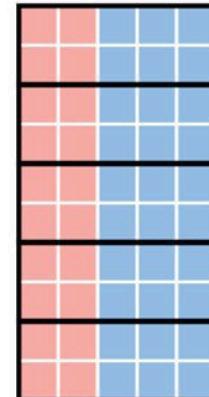
# 什么是 GERRYMANDERING, 不公平选区划分?

选区划分不公是指为了实现特定的政治结果而故意划分选区边界，比如将选民集中到较少的选区，或将选民分散到多个选区，以限制选民的权力

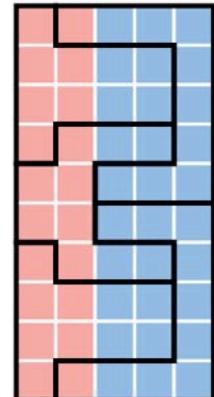
如何窃取选举



60% 蓝色  
40% 红色

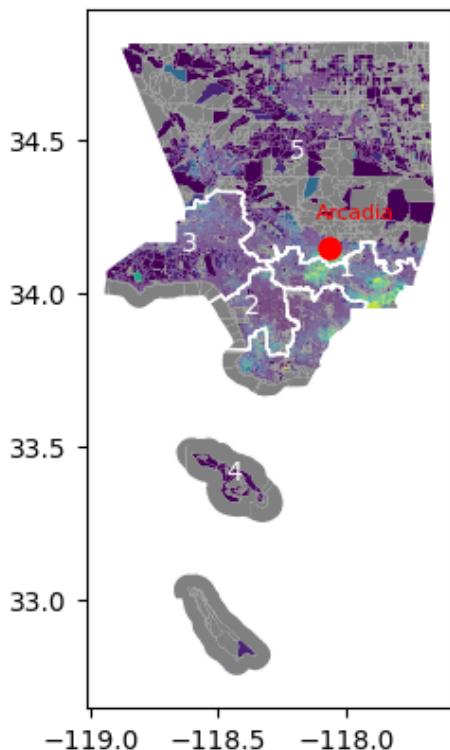


5 个区  
5 个蓝色  
蓝色获胜



5 个区  
3 个红色  
2 个蓝色  
红色获胜

洛杉矶县亚裔人口分布



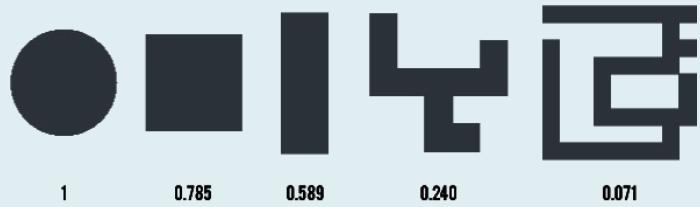
## 你为什么应该关心?

选区划分方式应代表社区。例如，洛杉矶县的亚裔人口为 14.7%，而阿卡迪亚的亚裔人口为 57%。您觉得您的政府领导人体现了这一点吗？选区通过代表和影响日常生活的问题赋予社区投票权，从地方层面到每 4 年一次的大选。

# 数字如何帮助识别不公正的选区划分？

## 紧凑

很多分数都是通过观察区域的形状来衡量紧凑度的。例如，Polsby-Popper 分数将区域与圆形进行比较，因此圆形区域比细长区域更好。但是，如果只看形状，当河流等自然地理条件是良好的边界线，但得分却很低时，就会出现问题。许多分数也没有考虑人口或人口。



紧凑性 Polsby-Popper 分数

## 政党公平

还有一些方法可以检测党派划分不公，即某一党派因选区而具有优势。但分数通常不足以完全衡量公平性，因为它们会假设选民的行为来做出预测。

**比例原则：**政党赢得的席位数量应与获得的选票百分比相匹配。

$50\% \text{ 的选票} = 50\% \text{ 的席位}$ 。然而，在竞争激烈的选区，对政党的偏爱发生微小变化可能会导致比例原则发生巨大变化。

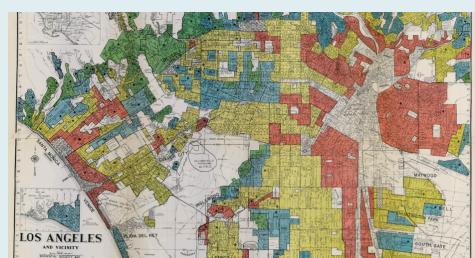
**党派对称：**如果各党派获得的选票比例相同，则应获得相同数量的席位。如果民主党获得 60% 的选票并获得 5/10 席位，那么在另一场选举中，如果共和党获得 60% 的选票，他们也应该获得 5/10 席位。

**低效率差距：**政党应该“浪费”相同数量的选票（赢得或输掉选举不需要的选票）。

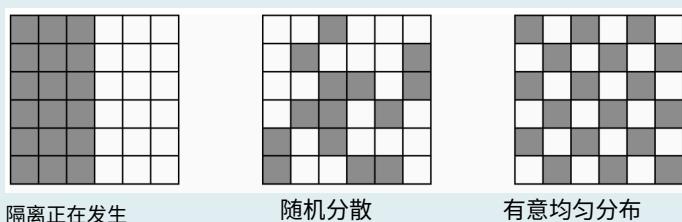
## 隔离

从历史上看，美国通过以下方式创造了种族隔离 **redlining, 红线**，一种不公平的方式，利用抵押贷款、住房贷款等手段将移民和黑人群体限制在特定地区和低收入人群中。在洛杉矶，红线政策造成了无家可归危机等后果，尤其是在 Highland Park, 和 Boyle Heights 等地区。

**Moran's I** 是通过检查一个地区的人口分离、均匀混合或聚集程度来衡量隔离程度的分数。



1939 年洛杉矶种族隔离和红线地图



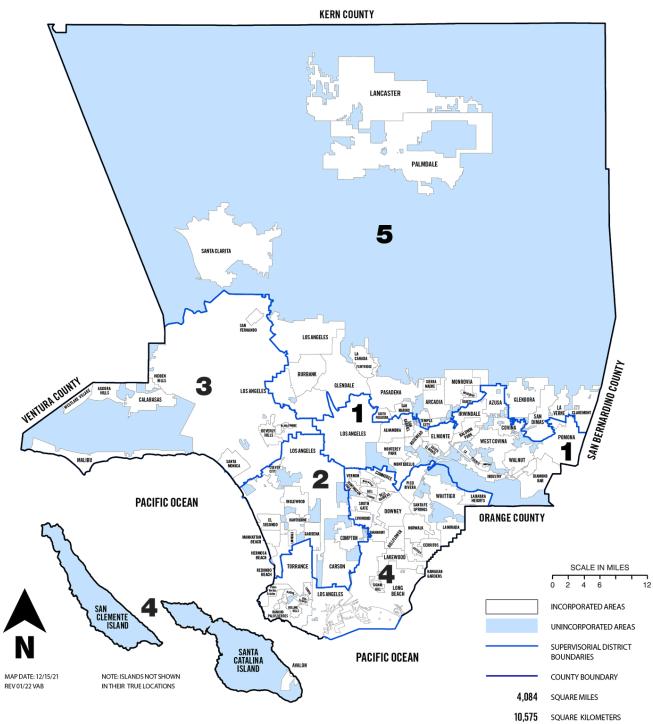
隔离正在发生

随机分散

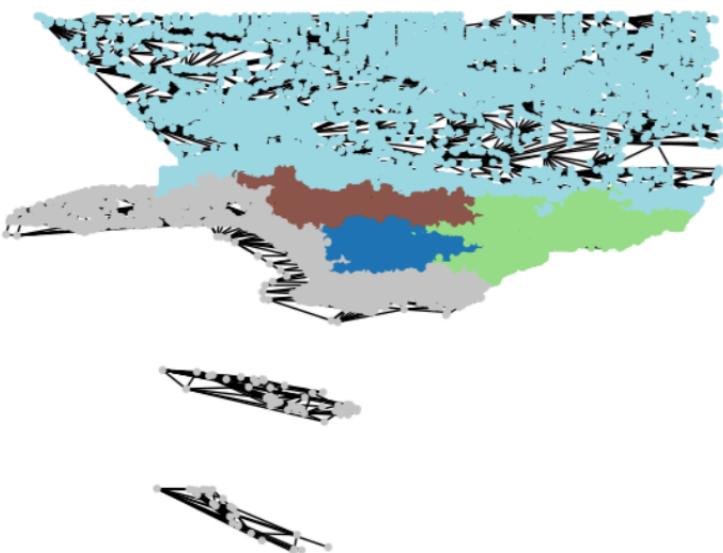
有意均匀分布

# 那么使用电脑怎么样？

在划分选区时，地图会变成“图表”来表示投票区如何连接，并帮助规划选区如何分组。图表由点和连接线组成，其中点代表城镇或选区，线表示哪些选区相邻且相邻。



洛杉矶县地图变成了包含不同区域的随机抽样图表



因为有这么多不同的可能方案，所以需要有方法来创建和比较大量不同的方案。“**随机抽样**”是一种随机挑选大量可能的地图并查看它们是否公平或平衡的方法。“**随机游走**”是一种在图表上以随机方向逐步游走的方法，用于分区，通过进行微小更改来探索划分区域的不同方法。“**马尔可夫链**”使用这些随机游走，根据概率和既定规则（如人口数量或紧凑性）从一个方案转到另一个方案。

# 那么使用电脑怎么样？

**ReCom**, 或重组，使用随机游走和马尔可夫链来创建新的区域，具体方法是将两个区域合并并随机拆分以制定新计划。在游走的每一步。

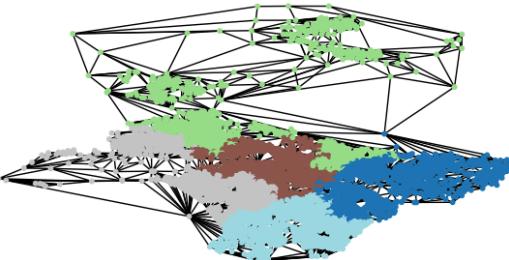
**Ensemble Analysis**, 集成分析是建立一组计划并将其与当前计划进行比较的一般过程。

**Gerrychain** 是代码中包含“ReCom”的特定工具，我使用它与集成分析来分析洛杉矶县的各个区域。

1. 我将一张包含人口、种族、投票历史和县区数据的洛杉矶县地图变成了一张“图表”。
2. 我制定了一个“随机”的选区划分计划
3. 我根据我的随机计划运行了 10,000 和 20,000 个“步骤”的 ReCom。
4. 我使用“集成分析”将我生成的计划中的平均亚裔人口和拜登的多数总统选票与 2021 年的计划进行比较。
5. 我还将 10,000 步和 20,000 步进行了比较，以确保我运行了足够的步骤

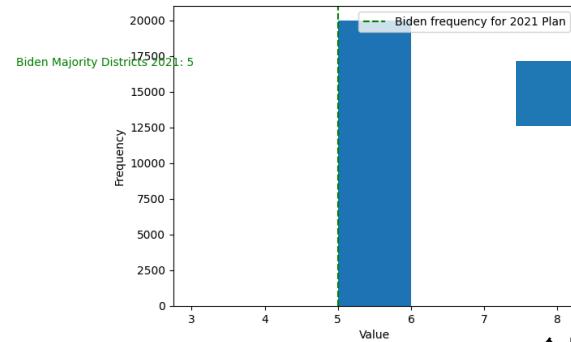
我发现生成的计划和 2021 年的计划都让所有 5 个选区的大多数选民投票给拜登。他们的亚裔人口数量相似，但没有一个选区的亚裔占多数。由于生成的计划和 2021 年的计划相似，因此不存在极端的党派或种族划分。

洛杉矶县初步重新划分选区计划



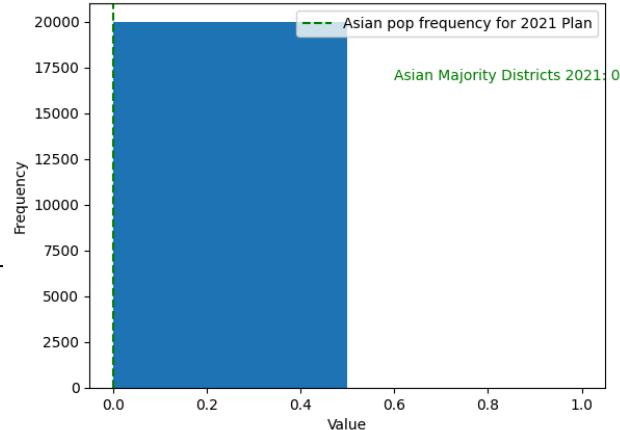
步骤 1 和 2

拜登 多数 总统 选票 地区 20K 步骤

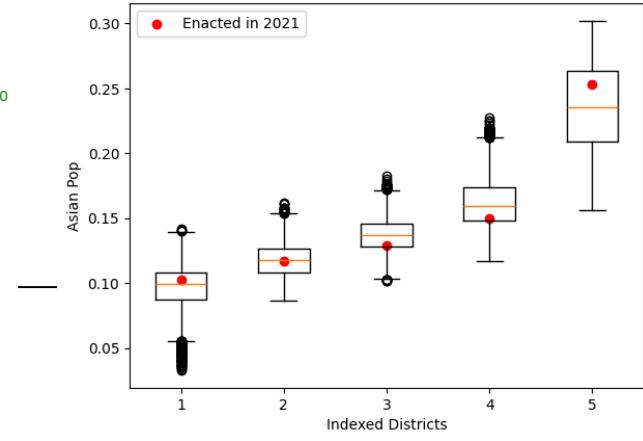


步骤4，集成分析

亚裔占多数地区的直方图 20k 步



5 个地区的亚裔人口 20K 步



# 挑战

在使用数学或计算机解决重新划分选区问题时，会面临很多挑战。

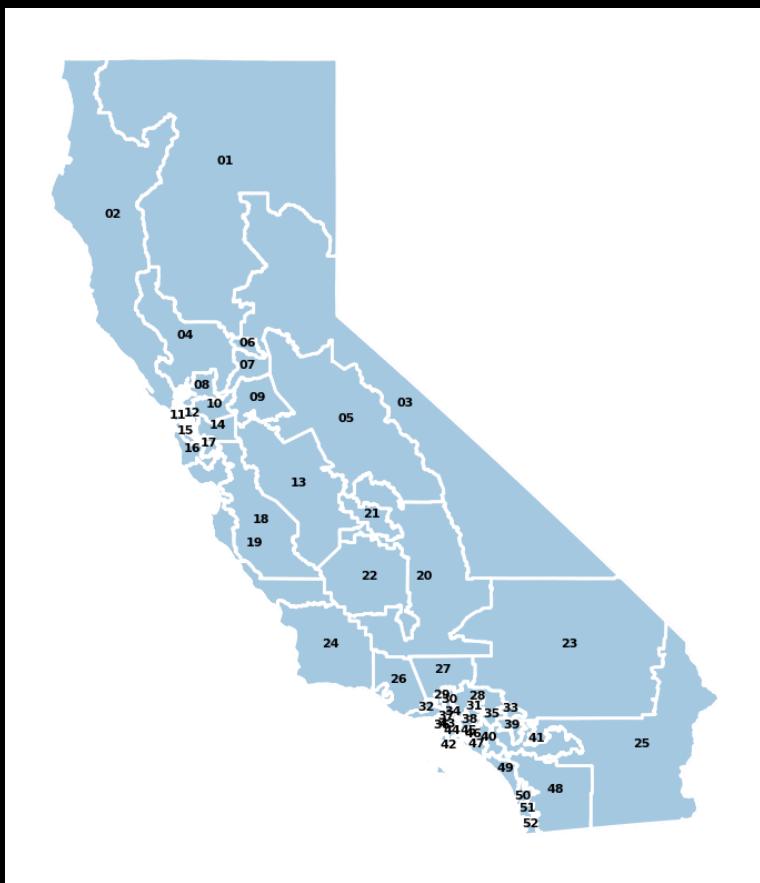
1. 获取准确、标准化的数据非常困难，尤其是有关人员或地理的数据。数据可能对政治、边界、水等自然地理等敏感。
2. 在衡量公平性或寻找选区划分不公时，数字和计算机是一个很好的起点，但不应是唯一的解决方案。数字和计算机也可能存在偏见和局限性，因为它们并不总是考虑重要的历史或人物——而这些总是在各个地方发生变化。

# 谢谢！

azhou4013@scrippscollege.edu

代码:

技术附录和来源:



## Technical Appendix

Data Sources for Code:

- [https://public.gis.lacounty.gov/public/rest/services/LACounty\\_Dynamic/Demographics/MapServer/16](https://public.gis.lacounty.gov/public/rest/services/LACounty_Dynamic/Demographics/MapServer/16)
- [data.census.gov](http://data.census.gov)
- <https://redistrictingdatahub.org/dataset/california-2020-general-election-results-disaggregated-to-the-2020-block/>
- [https://www.arcadiaca.gov/discover/city\\_manager\\_s\\_office/city\\_clerk/redistricting.php](https://www.arcadiaca.gov/discover/city_manager_s_office/city_clerk/redistricting.php)
- <https://gis.data.ca.gov/datasets/CDEGIS::us-congressional-districts/explore?location=37.593681%2C-116.251962%2C6.33>

Background:

For my final project, I decided to create an informational booklet introducing concepts from math of political districting at a high level aimed towards people like my parents in the LA area. Specifically, I created a booklet in Chinese aimed towards an audience of Chinese speaking immigrants who don't have a very technical background. I chose to go over redistricting at the congressional, county, and city level to give my audience a better idea of how redistricting affects them at a country and local level.

I first created visualizations for California's congressional districts and went over how redistricting worked to give more context to my audience. To give further context, California is divided into 52 Congressional districts, with LA County being part of the 28th Congressional district. Every 10 years, after the federal government updates the census, California redraws its districts with an independent commission created by voters in 2008. The commission includes 5 Republicans, 5 Democrats, and 4 with no party affiliation. Ranked in order of importance, the commission creates districts based on the following criteria: equal population, Voting Rights Act, contiguity, maintaining communities of interest, compactness, and nesting. The commission also considers public input, such as drawing maps, submitting comments, calling in, or joining advocacy groups. In the last redistricting round in 2021, California lost a congressional seat for the first time ever, due to slower population growth, pandemic, and an undercount due to a citizenship question on the census.

I then focused my ensemble analysis on LA County and Arcadia because I grew up in Arcadia, a city part of a predominantly high Asian population area within LA County. LA

County itself is 14.7% Asian, with the Asian population concentrated in areas like the San Gabriel Valley, central L.A. and South Bay. Arcadia is 57% Asian. Knowing this, I first created visualizations that showed the Asian population fraction and district boundaries for LA County and Arcadia. I also created a visualization for LA County that showed Trump vs Biden dominance to show partisan distribution. I then gave further context for how redistricting works for the county and for Arcadia.

In LA County, the current plan was drawn by the independent Citizens Redistricting Commission made up of all LA County residents for the first time in 2021 following the same rules as congressional districts. Before this, districts were drawn individually by the supervisors for their own districts. The 5 districts represent LA County's 88 cities and communities in LA County Supervisor elections. The commission's goal was to give Latino and Asian residents better representation while maintaining a concentration of Black voters. There are some critiques that adding predominantly white communities to the 2nd District, a district that has 29% Black voters, might dilute Black votes. The map also added a second majority-Latino district and increased the 1st District's Asian voter percentage to 27%. An interesting note is that LA County has never had an Asian American supervisor because Asian voters are concentrated in areas that are widely separated geographically. In LA County, the islands (San Catalina and San Clemente) are connected to the 4th district, the closest district by the coast. In Arcadia, the city is divided into 5 districts, each with an elected City Council Member. With public input, the City Council redraws districts every 10 years following the same rules as above and while following natural boundaries like rivers, streets, etc. and without partisan gerrymandering.

To create my visualizations and run ensemble analysis, I primarily used the gerrychain library on python. For California's congressional districts, I just plotted a shapefile from the California State Geoportal on GIS, which directly received the specific boundary data from the independent citizens redistricting commission, aggregated by the California Department of Technology. For LA County and Arcadia, I first joined together two shapefiles. The first was Census Blocks 2020 customized for LA County with data from the Census Bureau, including the population in 2020 and 2021's created county supervisor districts. The second was California 2020 General Election Results Disaggregated to the 2020 Block from Redistricting Data Hub. The block shapefiles are originally from the Census Bureau's Public Law 94-171 dataset and were combined with election results from the Harvard Voting and Election Science Team. I

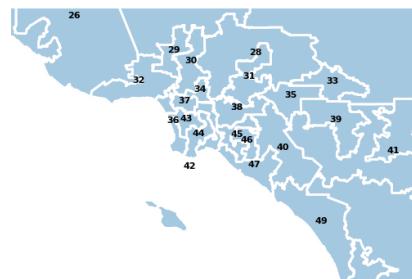
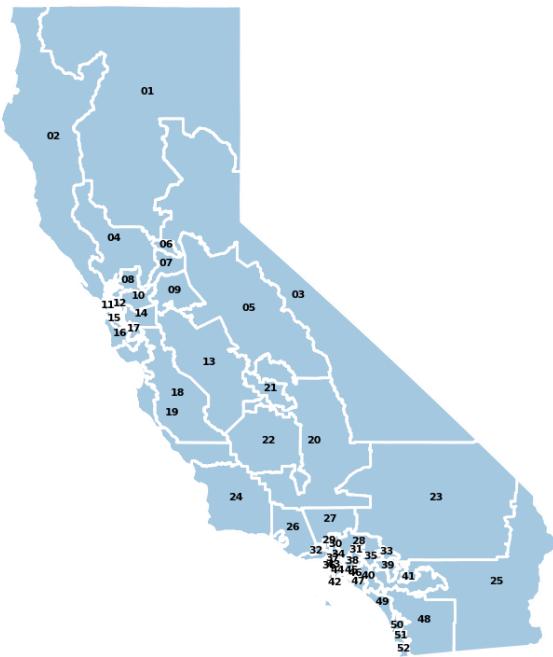
joined these together through Census Tract and Census Block. After joining, I then imported additional data about race from the US Census by filtering for LA County block tract 2020 P1. I joined again by Census Tract and Census Block. Then, I was able to calculate the Asian population fraction and voting distribution. For Arcadia, I filtered this data frame by city to only look at Arcadia's Asian population. I then plotted city council boundaries on top of this map with data from Arcadia's government website, which maintained a GIS map that I converted to a SHP file.

For ensemble analysis, I ran the Markov Chain for 20000 steps from a random initial districting plan. The first initial districting plan was created at the precinct level and took too long to run, so I dissolved and grouped by block group instead for my initial districting plan. I looked at cutedges over time, Asian population majority and Asian population, and Democrat majority. I chose Asian population majority and Asian population because of my audience and LA County's demographics. I used presidential votes for Democrat versus Republican because California is a majority Democratic state. I then looked at my Markov Chain at 10000 steps versus 20000 steps to make sure my chain had run for enough steps. I then compared my results to 2021's enacted county plan.

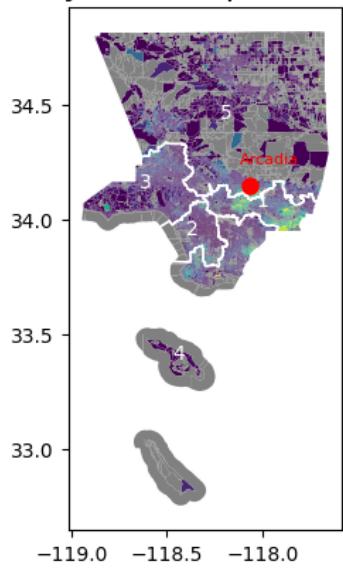
#### Analysis:

I believe my Markov chain has run for enough steps and has reached close to stationary distribution because when comparing my histograms and boxplots, 10k steps and 20k steps are very similar. Each had 0 districts with a majority Asian population, and 5 Democrat majority districts, which means that running for 10k steps was enough. According to my boxplots, the Asian population of 2021's plan is within the range of my ensemble's Asian population. This suggests that there isn't intense racial or partisan gerrymandering, at least in terms of the plans explored by my ensemble analysis. However, my analysis is still limited to the accuracy of my data sources and using outcomes of 2020's presidential election to determine partisanship. Next steps can include looking at other demographics, like LA county's significant Hispanic population, applying county level election data, or 2024's election.

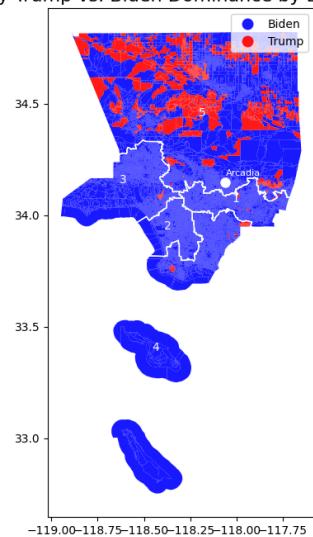
## Figures:



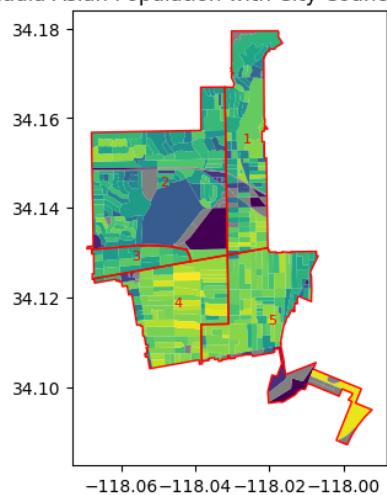
LA County Asian Population Fraction



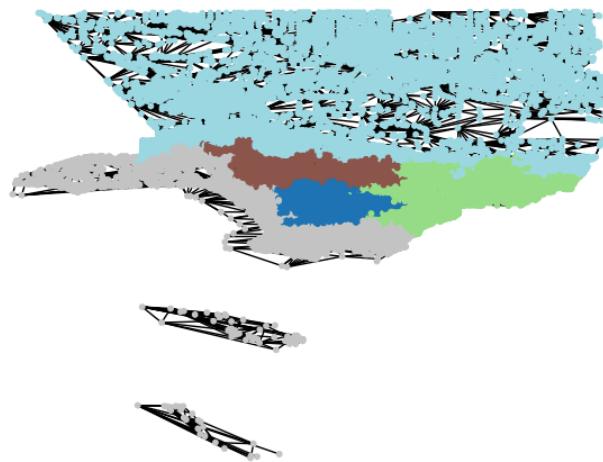
LA County Trump vs. Biden Dominance by District in 2020



Arcadia Asian Population with City Council Districts

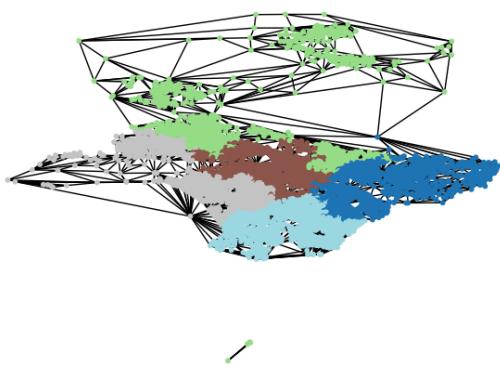


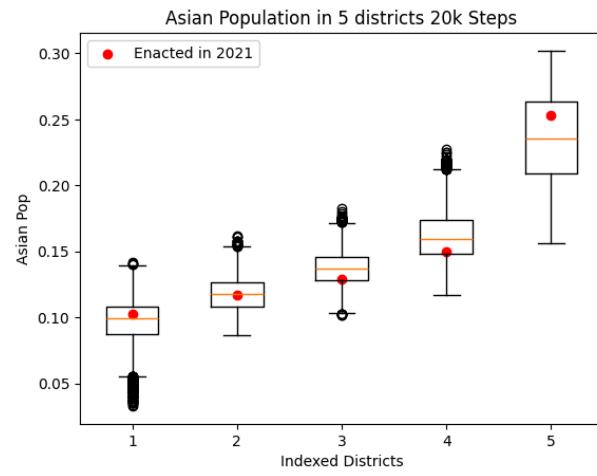
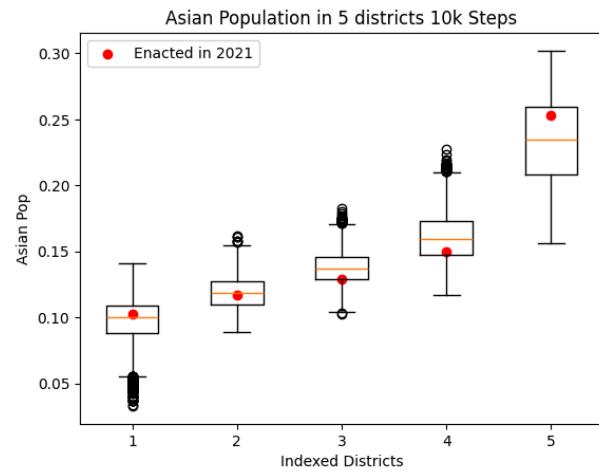
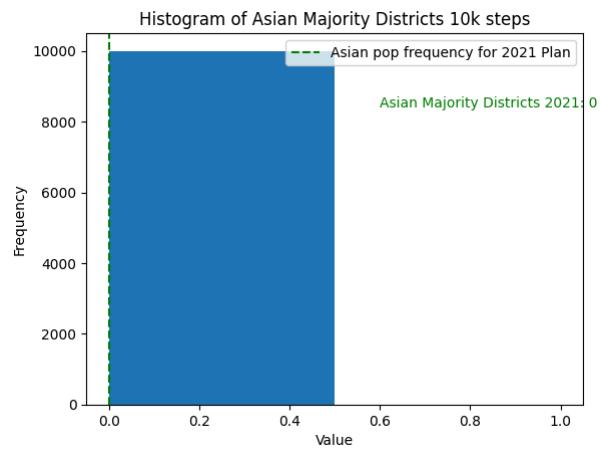
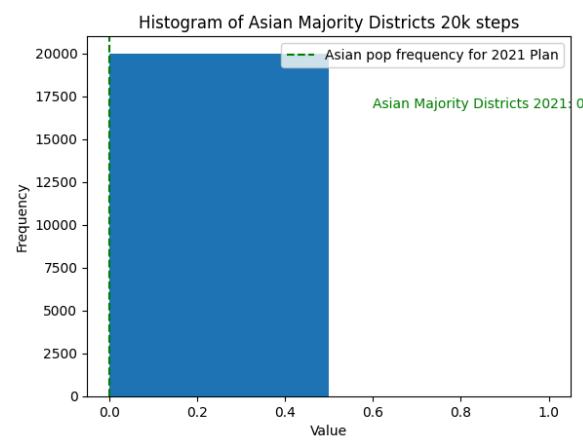
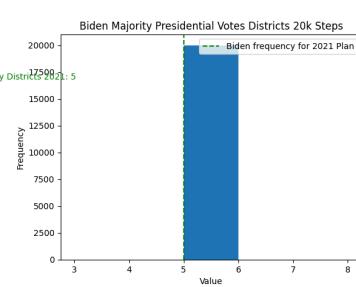
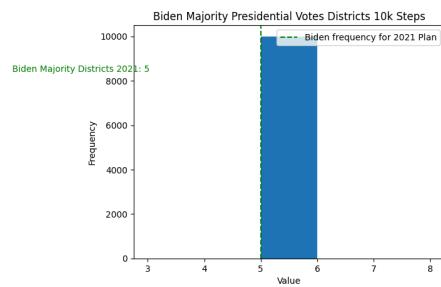
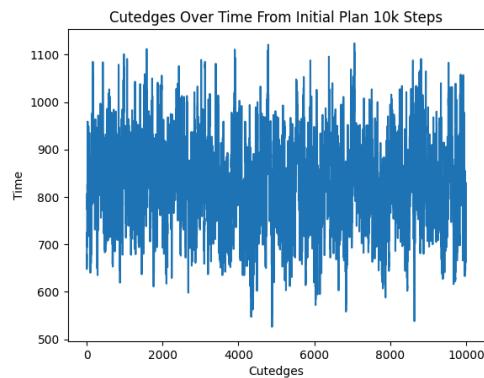
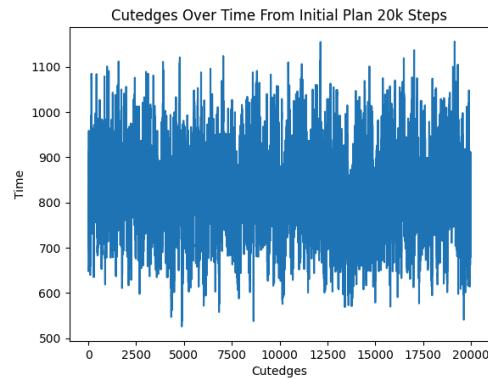
LA County Initial Districting Plan



Dissolved/grouped by block group:

LA County Initial Districting Plan





Other references:

- [Arcadia's Voting Districts Are Changing: What's That Mean for You? | Arcadia, CA Patch](https://www.arcadiapatch.com/2021/12/15/arcadia-s-voting-districts-are-changing-whats-that-mean-for-you/)
- <https://www.latimes.com/california/story/2021-12-15/la-county-supervisors-redistricting-map-finalized>
- [https://redistricting.lacounty.gov/final\\_map\\_and\\_submissions/](https://redistricting.lacounty.gov/final_map_and_submissions/)
- <https://mqqq.org/gerrybook.html>
- <https://calmatters.org/politics/2021/11/california-redistricting-what-you-need-to-know/>
- <https://radnorite.com/6643/opinion/the-destructive-art-of-gerrymandering/>
- <https://lacounty.gov/government/about-la-county/redistricting/>
- [https://clkrep.lacity.org/onlinedocs/2019/19-0600\\_misc\\_5-6-19.pdf](https://clkrep.lacity.org/onlinedocs/2019/19-0600_misc_5-6-19.pdf)
- <https://hdsr.mitpress.mit.edu/pub/1ds8ptxu/release/5>