



# Coropletas y paletas

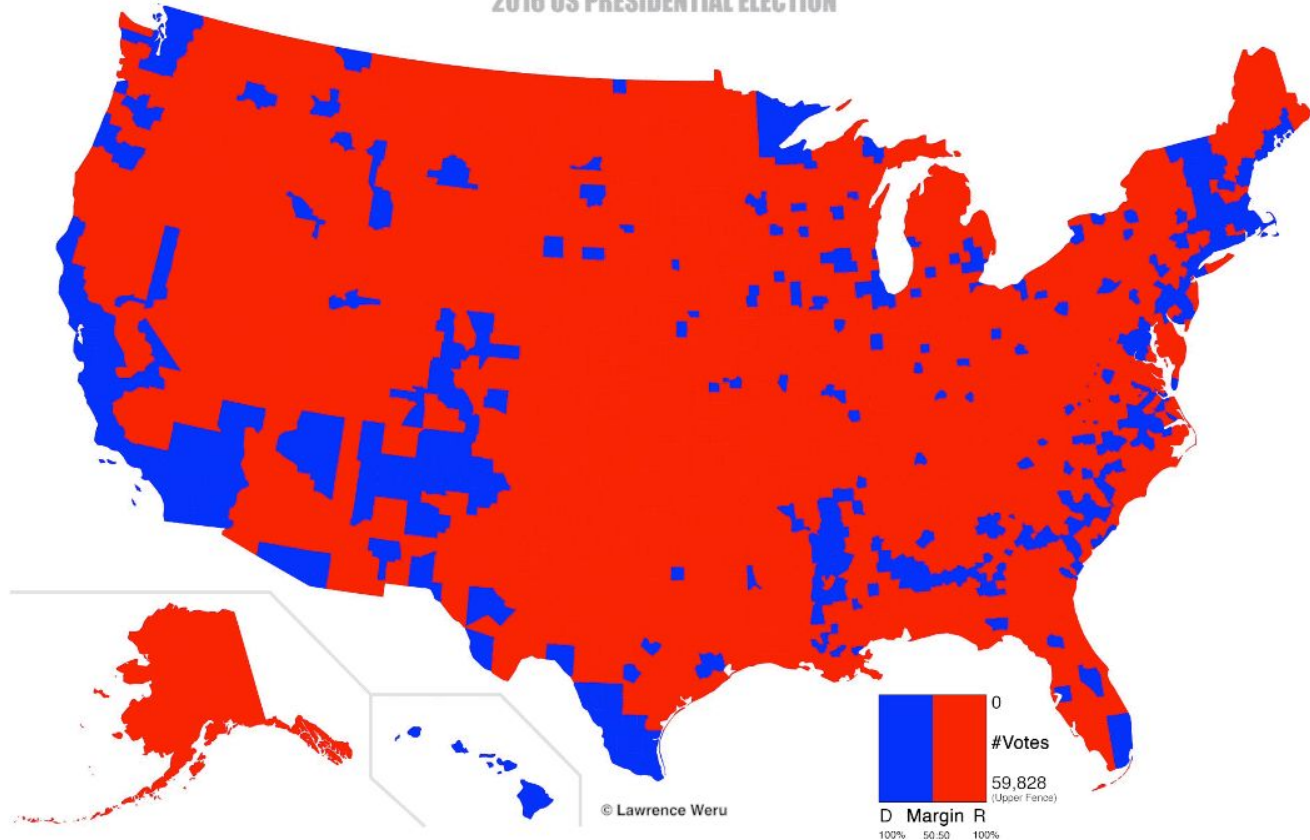
Mapeando muchas variables



# MUDDY AMERICA

VOTE MARGINS + VOTE TOTALS

2016 US PRESIDENTIAL ELECTION

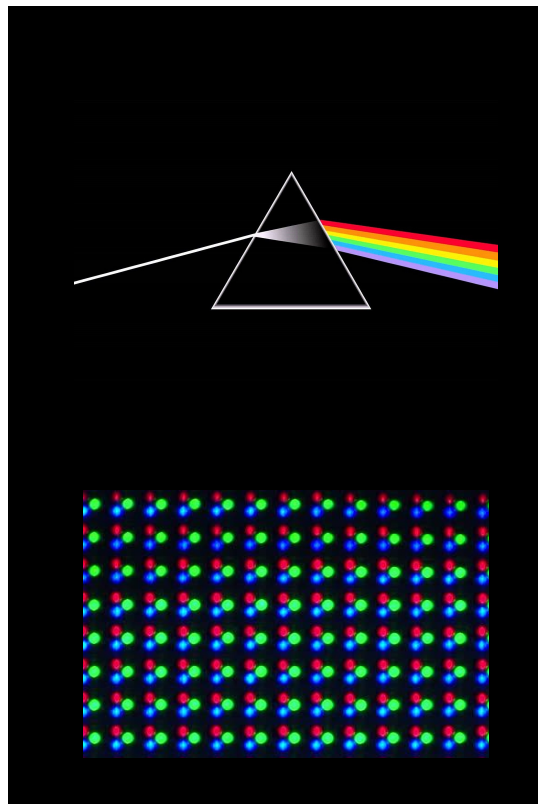


# Pero antes, un poco de teoría del color (luz)

Modelo matemático del color. Podemos descomponer la luz en dos tipos de componentes de acuerdo al modelo que usemos:

- HSL (Hue, Saturation, Lightness)
- RGB (Red, Green, Blue)

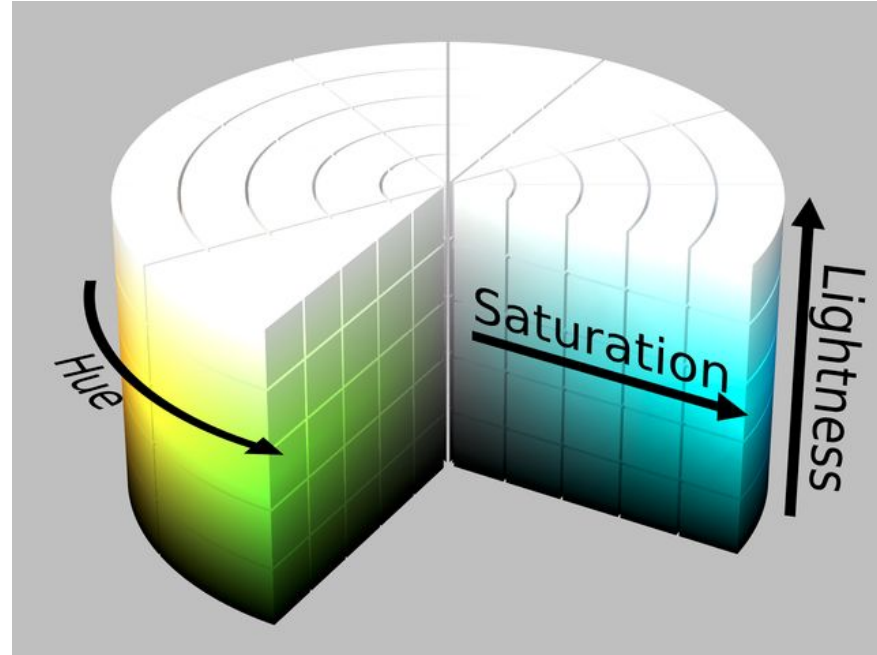
Para más recursos pueden ver [este](#) material y [este](#) de la clase introducción a las Ciencias de la Computación de Standord



# Modelo HSL

Cada componente del modelo tiene su unidad de medida (grados, porcentajes, enteros) que representa alguno de los elementos del modelo.

- *Hue* o Matiz: en grados de un círculo cromático
- Saturación en %
- Cantidad de luz o *Lightness* en %

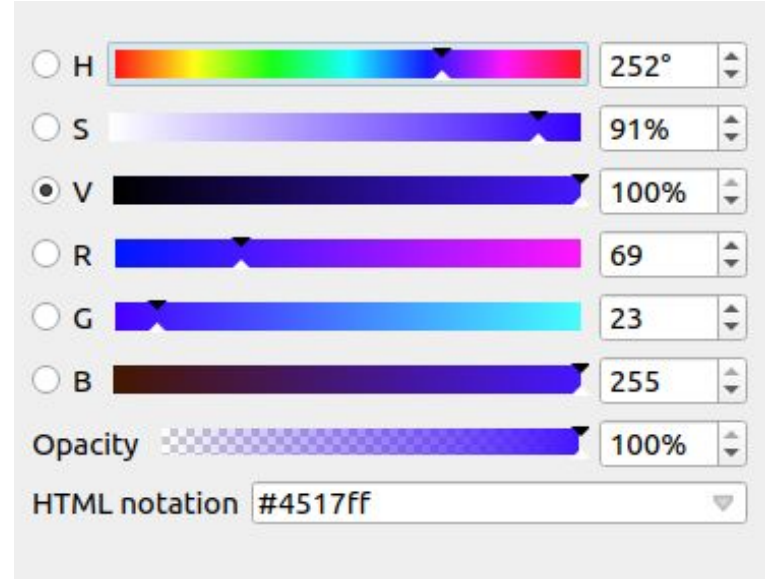


# Modelo RGB

- Cantidad de Rojo: de 0 a 255
- Cantidad de Verde (*Green*): de 0 a 255
- Cantidad de Rojo (*Red*): de 0 a 255

Puede haber otro espacio de 0 a 255 para la transparencia o *alpha* en ambos modelos

255 es el valor máximo que se puede alcanzar por un número binario de 8 dígitos o bits



A color picker interface with the following settings:

- H** (Hue): 252°
- S** (Saturation): 91%
- V** (Value): 100%
- R** (Red): 69
- G** (Green): 23
- B** (Blue): 255
- Opacity**: 100%
- HTML notation**: #4517ff

$$255 = 2^8 - 1 = \text{FF}_{16} = 11111111_2$$

# Notación HTML

Solo otro modo de representar el modelo RGB.

No utiliza enteros de 0 a 255.

Utiliza valores de 0 a 9 y luego de A a F. Esto se denomina sistema hexadecimal o hex



<b>0</b> <sub>hex</sub> = 0 <sub>dec</sub> = 0 <sub>oct</sub>	0	0	0	0
<b>1</b> <sub>hex</sub> = 1 <sub>dec</sub> = 1 <sub>oct</sub>	0	0	0	1
<b>2</b> <sub>hex</sub> = 2 <sub>dec</sub> = 2 <sub>oct</sub>	0	0	1	0
<b>3</b> <sub>hex</sub> = 3 <sub>dec</sub> = 3 <sub>oct</sub>	0	0	1	1
<b>4</b> <sub>hex</sub> = 4 <sub>dec</sub> = 4 <sub>oct</sub>	0	1	0	0
<b>5</b> <sub>hex</sub> = 5 <sub>dec</sub> = 5 <sub>oct</sub>	0	1	0	1
<b>6</b> <sub>hex</sub> = 6 <sub>dec</sub> = 6 <sub>oct</sub>	0	1	1	0
<b>7</b> <sub>hex</sub> = 7 <sub>dec</sub> = 7 <sub>oct</sub>	0	1	1	1
<b>8</b> <sub>hex</sub> = 8 <sub>dec</sub> = 10 <sub>oct</sub>	1	0	0	0
<b>9</b> <sub>hex</sub> = 9 <sub>dec</sub> = 11 <sub>oct</sub>	1	0	0	1
<b>A</b> <sub>hex</sub> = 10 <sub>dec</sub> = 12 <sub>oct</sub>	1	0	1	0
<b>B</b> <sub>hex</sub> = 11 <sub>dec</sub> = 13 <sub>oct</sub>	1	0	1	1
<b>C</b> <sub>hex</sub> = 12 <sub>dec</sub> = 14 <sub>oct</sub>	1	1	0	0
<b>D</b> <sub>hex</sub> = 13 <sub>dec</sub> = 15 <sub>oct</sub>	1	1	0	1
<b>E</b> <sub>hex</sub> = 14 <sub>dec</sub> = 16 <sub>oct</sub>	1	1	1	0
<b>F</b> <sub>hex</sub> = 15 <sub>dec</sub> = 17 <sub>oct</sub>	1	1	1	1

# Qué y cómo mapear

Representar en un mapa un proceso electoral complejo como el de Estados Unidos es un desafío. Muchas variables en juego:

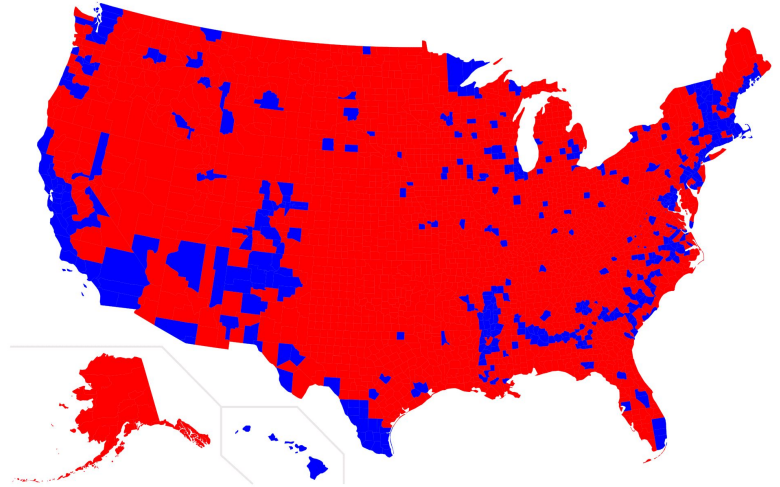
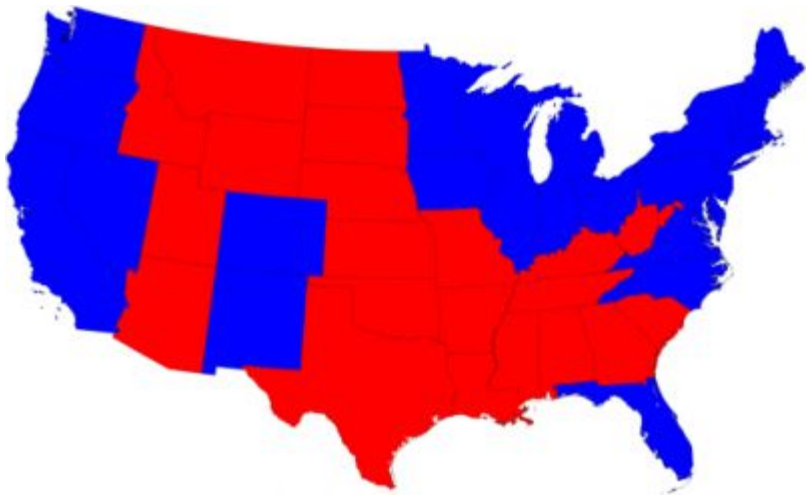
- Escala de la Unidad Espacial: estados, condados (PUEM o MAUP en inglés)
- Electores o votantes
- % de votos o n de electores

Veremos las decisiones que tomó [Larry Wedu](#) para crear Muddy America. Pueden leer luego este [post](#) o su [charla TED](#).

Para una guía de cómo funciona el proceso electoral estadounidense pueden seguir esta [serie de notas](#).

# Qué y cómo mapear

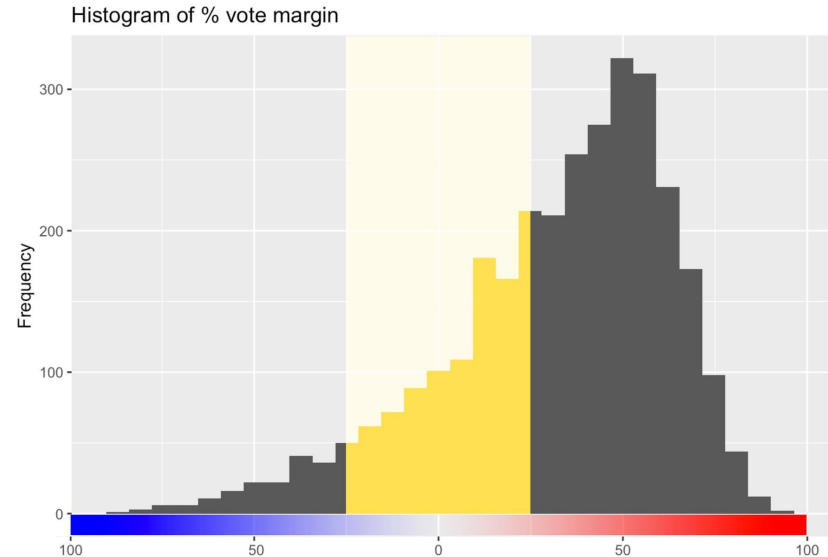
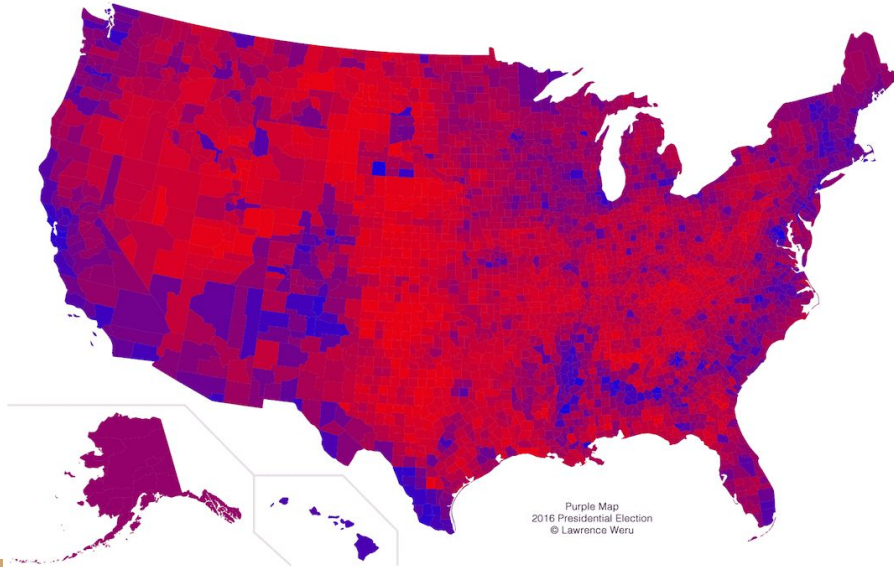
Mapas *Winner takes all*. Cualitativos y univariados. Hay variables que no se ven. Porcentaje marginales, cantidad de población





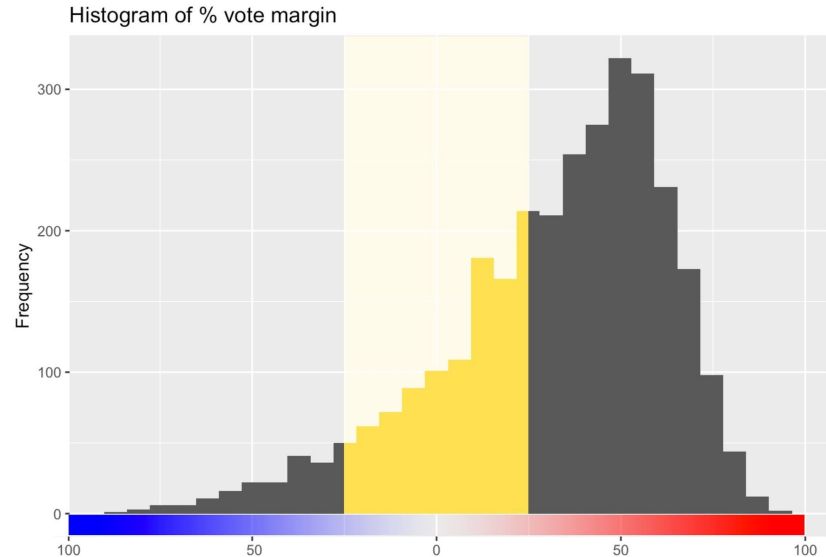
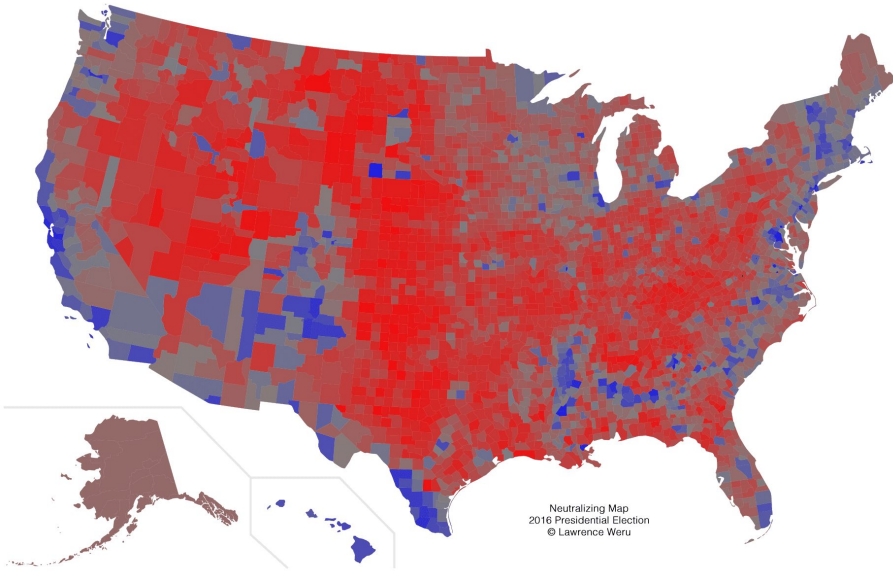
# Qué y cómo mapear

Este mapa incorpora el % de voto a cada partido (variable continua) con dos escalas de colores (Rojo y Azul) con un significado particular. Pero al ojo le cuesta cuantificar esa diferencia de violetas y los colores circundantes influyen en el color que percibe el ojo.



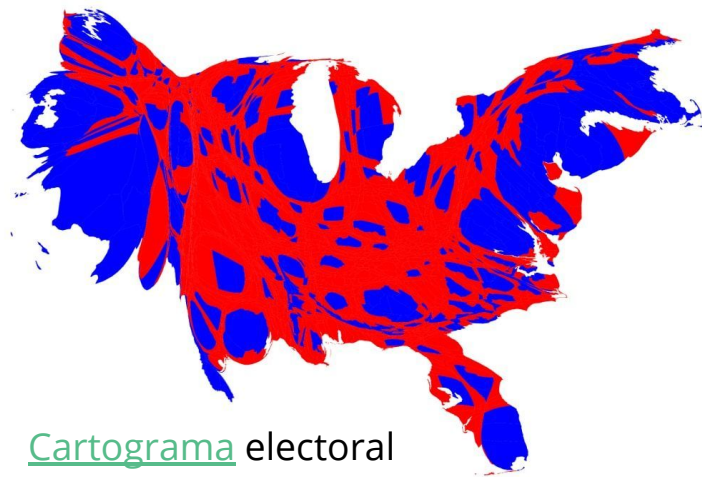
# Qué y cómo mapear

Convirtió ese rango de ambigüedad en una escala gris. Seguimos observando una única variable (% de distribución del voto en 2 partidos) en una dimensión continua. Pero se quiere visualizar otra: cantidad de población.

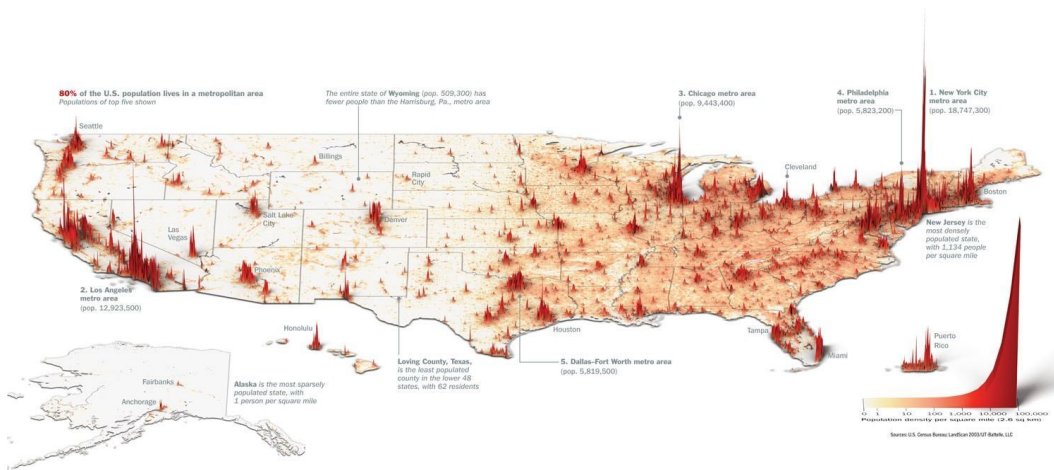
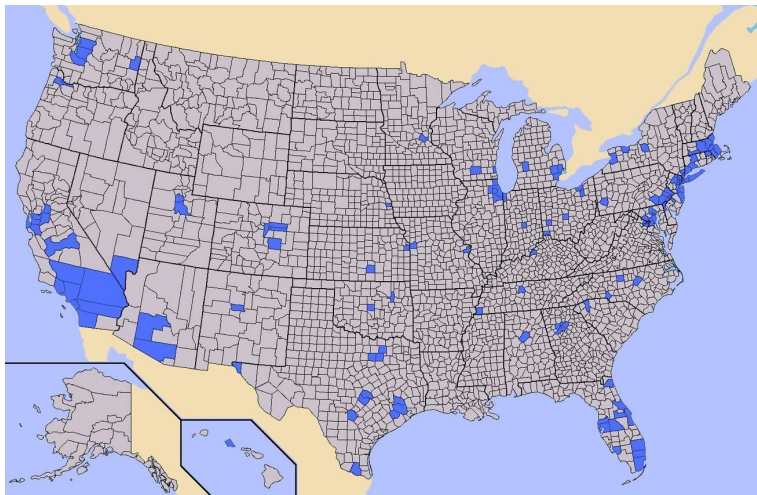


# Qué y cómo mapear

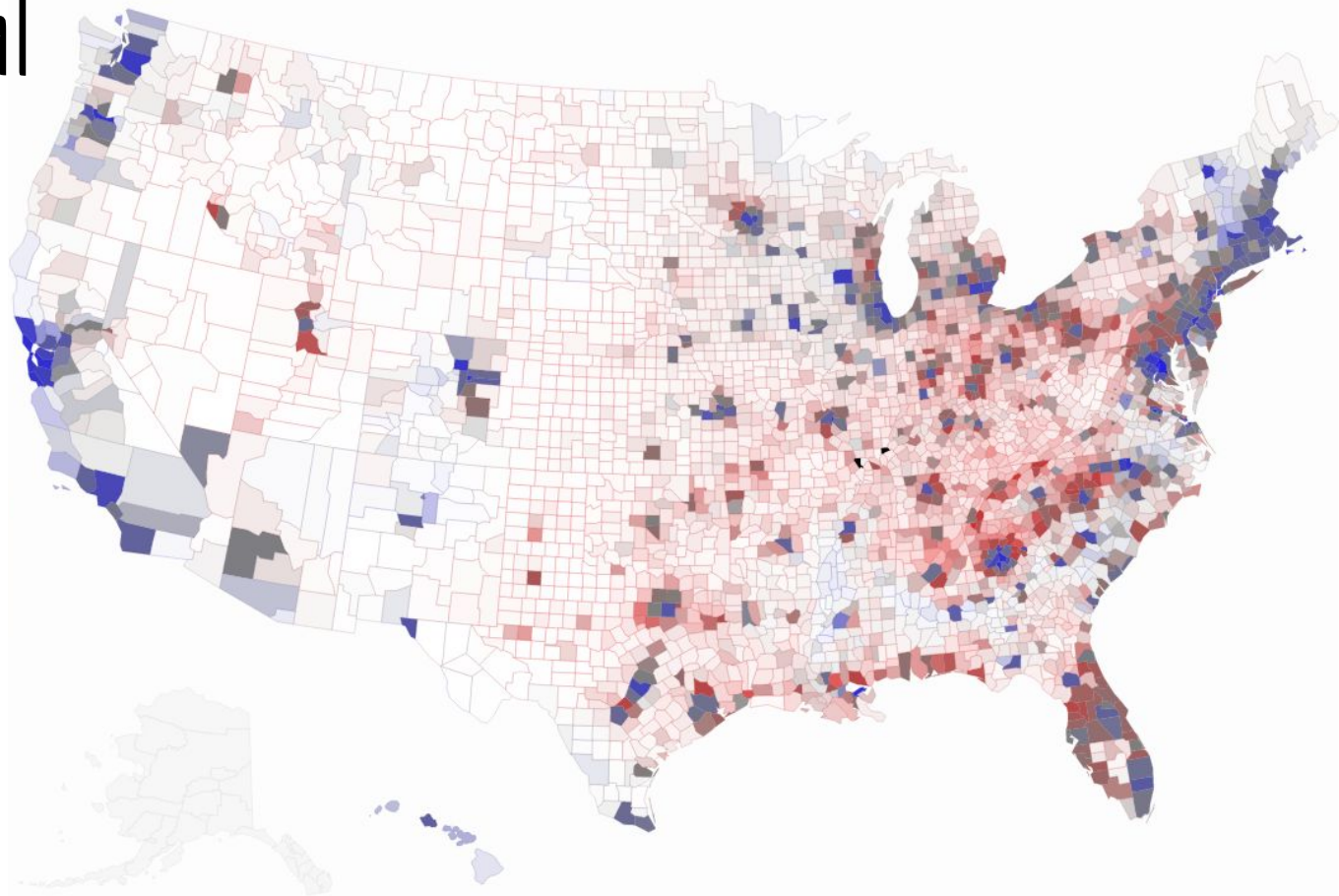
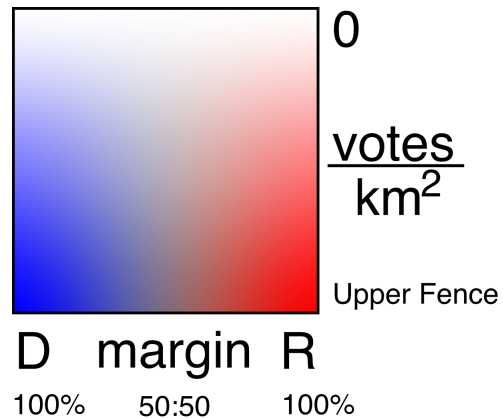
No todos los condados tienen la misma cantidad de votos emitidos ni la misma cantidad de electores.



Cartograma electoral



# Resultado final



# La fórmula de la paleta de color en modelo HSL

Esta paleta toma una representación en la siguiente fórmula

$$MuddyColor = HSL \left( Hue(winner(D, R)), \frac{|D - R|}{totalVote}, \frac{\left(1 - \frac{totalVote}{upperFence}\right) * 100}{2} + 50 \right)$$

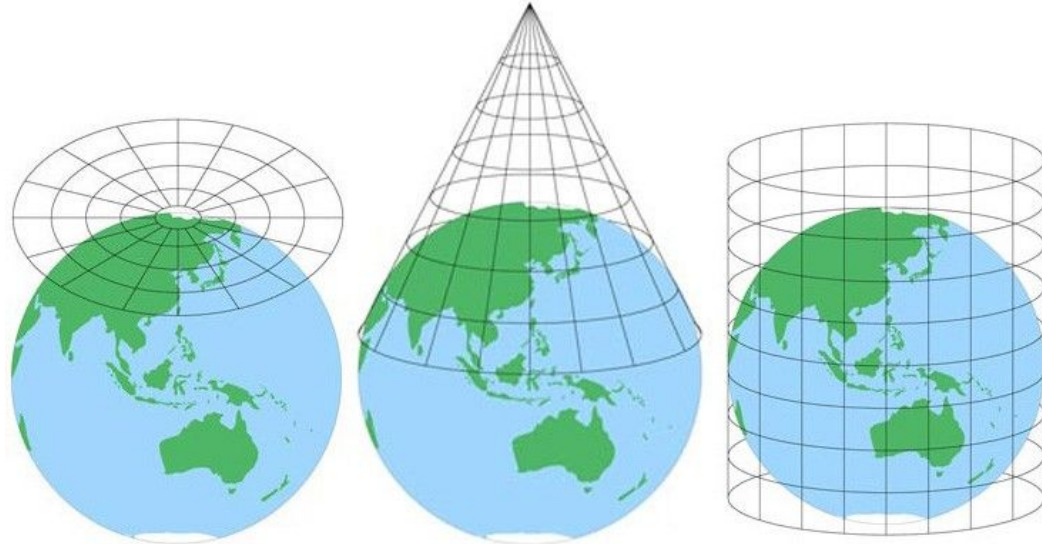
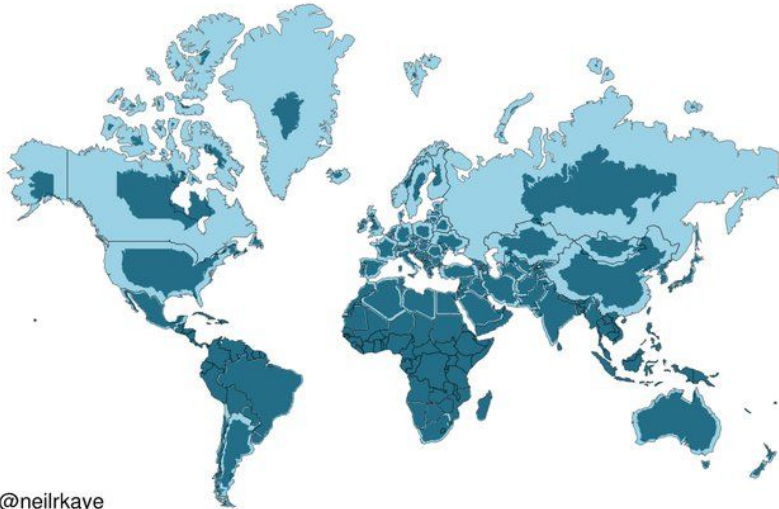




# Distorsión del espacio

Las proyecciones a utilizar para visualizar (puede ser distinta a la utilizada para procesar los datos) influye en el efecto distorsivo de la visualización para forma, distancia, área.

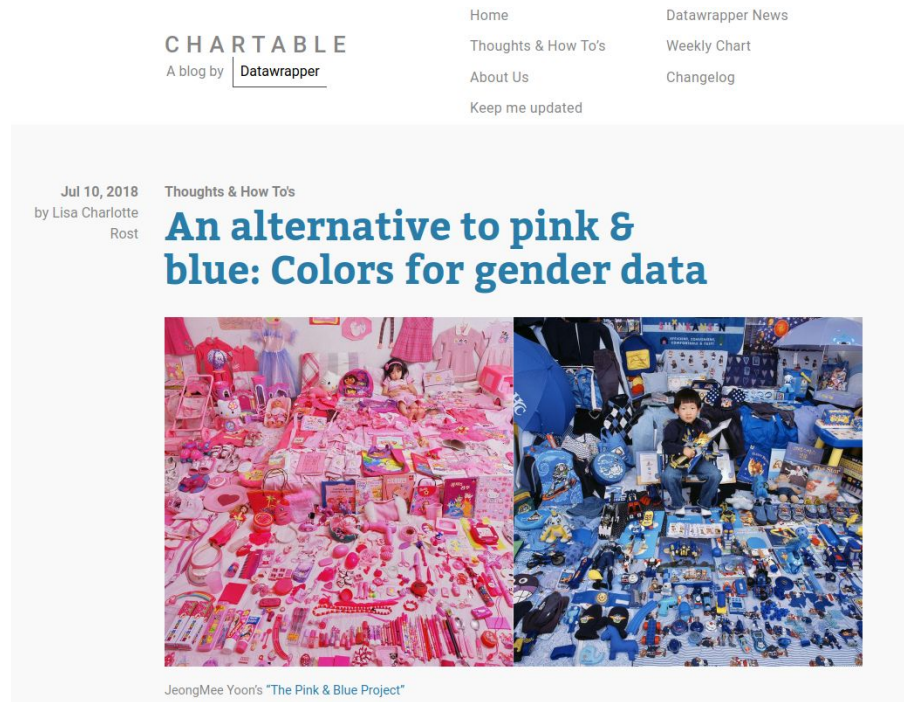
World Mercator projection with true country size added



# Las paletas también comunican

Las paletas no son solo estéticas.  
También comunican información.  
Puede ser positivo como el caso del  
rojo y azul en los partidos de EEUU.

Puede ser de modo negativo  
estigmatizando (utilizar rojo con sus  
connotaciones en mapas de pobreza,  
distribución del ingreso, etc) o el caso  
de usar rosa y celeste para visualizar  
por sexo y como puede reproducir  
estereoripos nocivos



# Las paletas también comunican

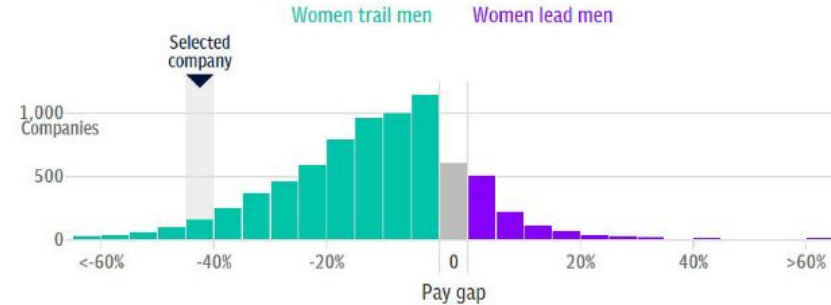
## Paleta de The Telegraph (RU)

Men	Women	Economist, 2018
Men	Women	Morgenpost 2017
Men	Women	Guardian, 2018
Men	Women	Financial Times, 2018
Men	Women	Telegraph, 2018
Men	Women	Information is beautiful, 2014
Men	Women	Washington Post, 2017
Men	Women	Bloomberg, 2018
Men	Women	BBC, 2017
Men	Women	BBC, 2018



### The national picture

Median hourly gender pay gap distribution across all companies

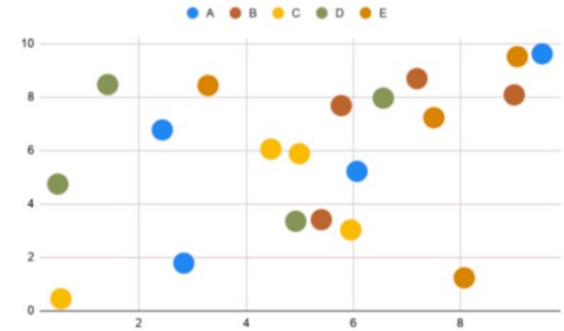
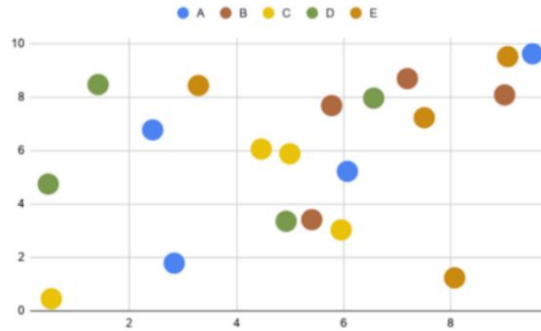




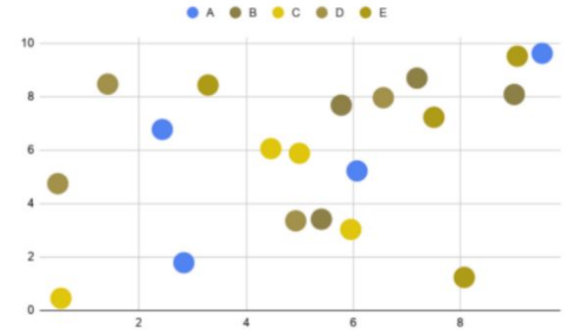
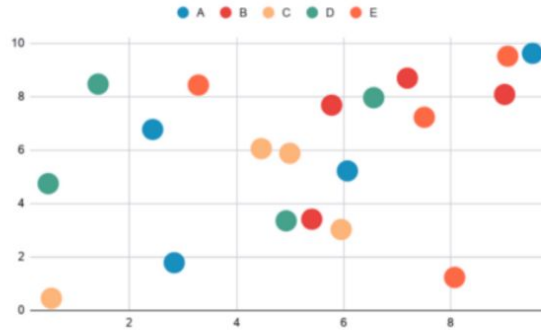
# El daltonismo es más común de lo que pensamos

El 5% de la población tiene daltonismo.

En una audiencia de 20 personas, es probable que al menos 1 se encuentre con dificultades para distinguir colores en la visualización si no se usan paletas acordes



First: Protanomaly (red-weak). Second: Deuteranomaly (green-weak). Created with Google Sheets and Cobliis.



First: Tritanomaly (blue-weak). Second: Protanopia (red-blind). Created with Google Sheets and Cobliis.

# Problema de la Unidad Espacial Modificable

Cuando trabajamos con estadísticos, casi siempre los calculamos a lo largo de un índice  $i$ . En el análisis espacial dicha  $i$  se construye a partir de un recorte arbitrario del espacio ( que puede ser recortado, literalmente, infinitas formas). La forma de recortar concreta que elijamos incide en nuestro análisis. En inglés se lo llama MAUP.

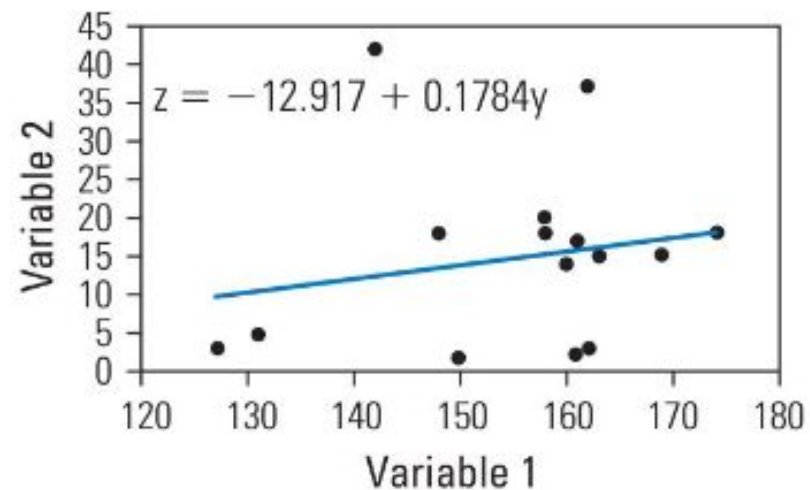
$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Variable 1

174	162	169	161
150	144	161	160
131	127	162	163
142	148	158	158

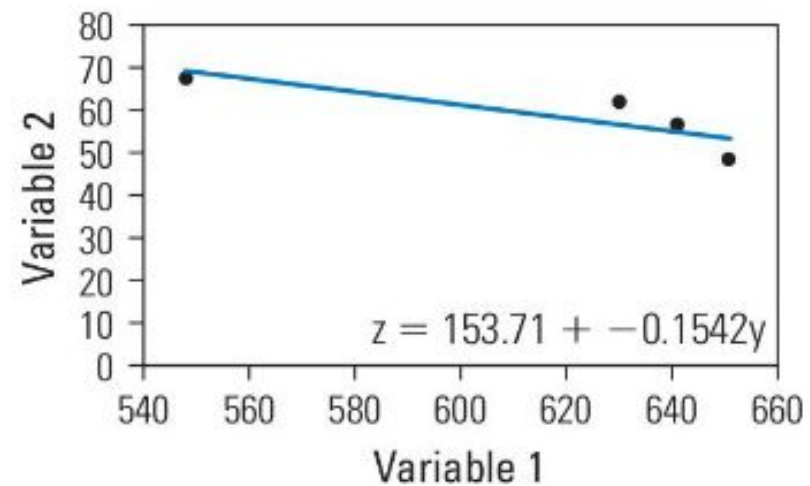
Variable 2

18	37	15	17
2	5	2	14
5	3	3	15
42	18	18	20



630	651
548	641

62	48
68	56



# PUEM

El **efecto de escala**: El análisis estadístico basado en datos agregados en áreas de diferentes tamaños producirá resultados diferentes

El **efecto de zonificación**: Dos zonas de igual área y forma, pueden producir diferentes resultados de acuerdo a su ubicación

Efecto escala

630	651
548	641

174	162	169	161
150	144	161	160
131	127	162	163
142	148	158	158

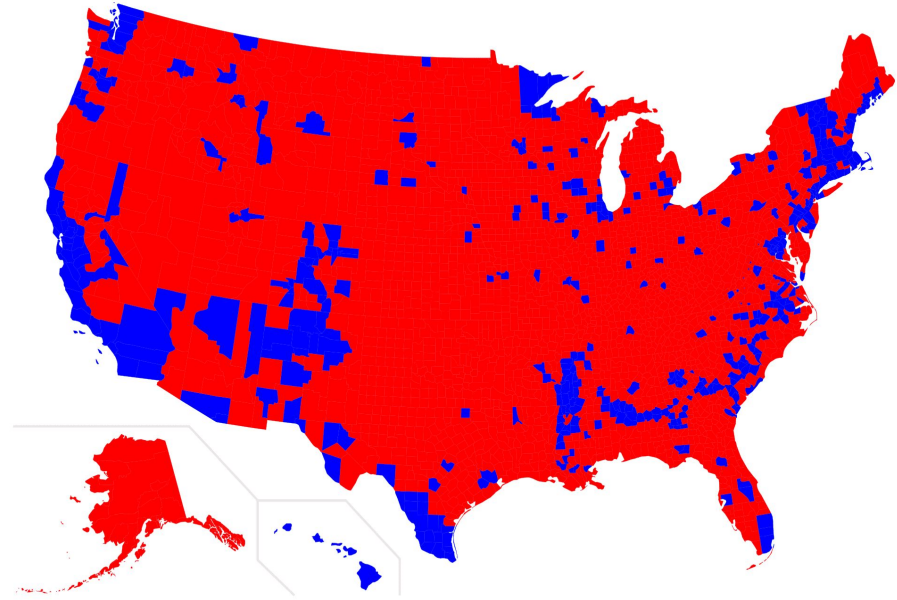
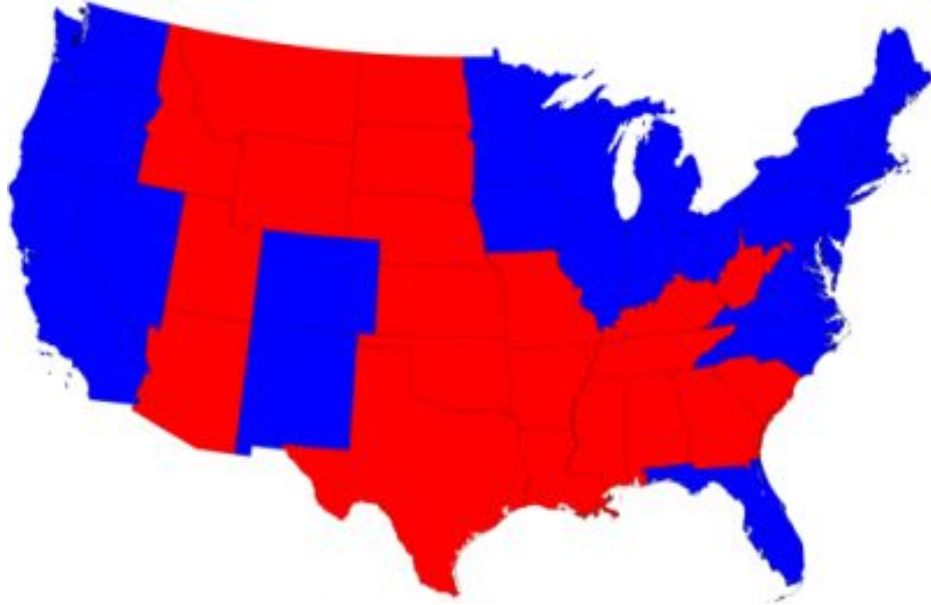
Efecto zonificación

968	972
530	

965	
526	979

PUEM

## El efecto de escala: resultados por estado o condado



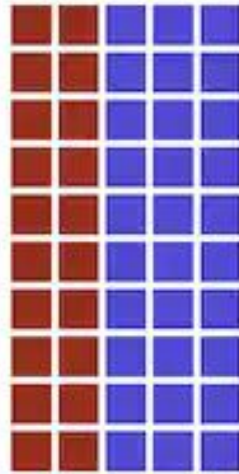
# PUEM

El efecto de  
zonificación:  
construcción de  
circunscripciones  
electorales.

[Gerrymandering](#).

Three different ways to divide 50 people into five districts

50  
people



60% blue,  
40% red

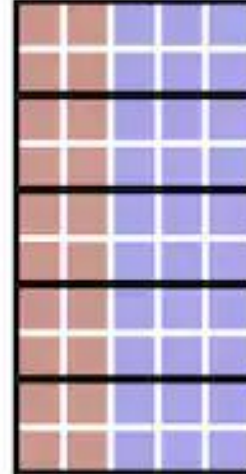
1. Perfect  
representation



3 blue districts,  
2 red districts

BLUE WINS

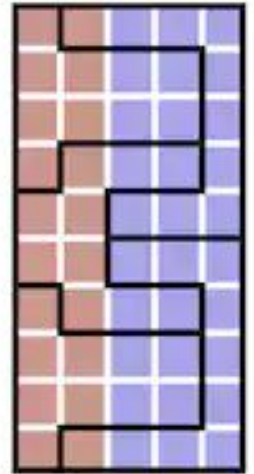
2. Compact,  
but unfair



5 blue districts,  
0 red districts

BLUE WINS

3. Neither compact  
nor fair



2 blue districts,  
3 red districts

RED WINS