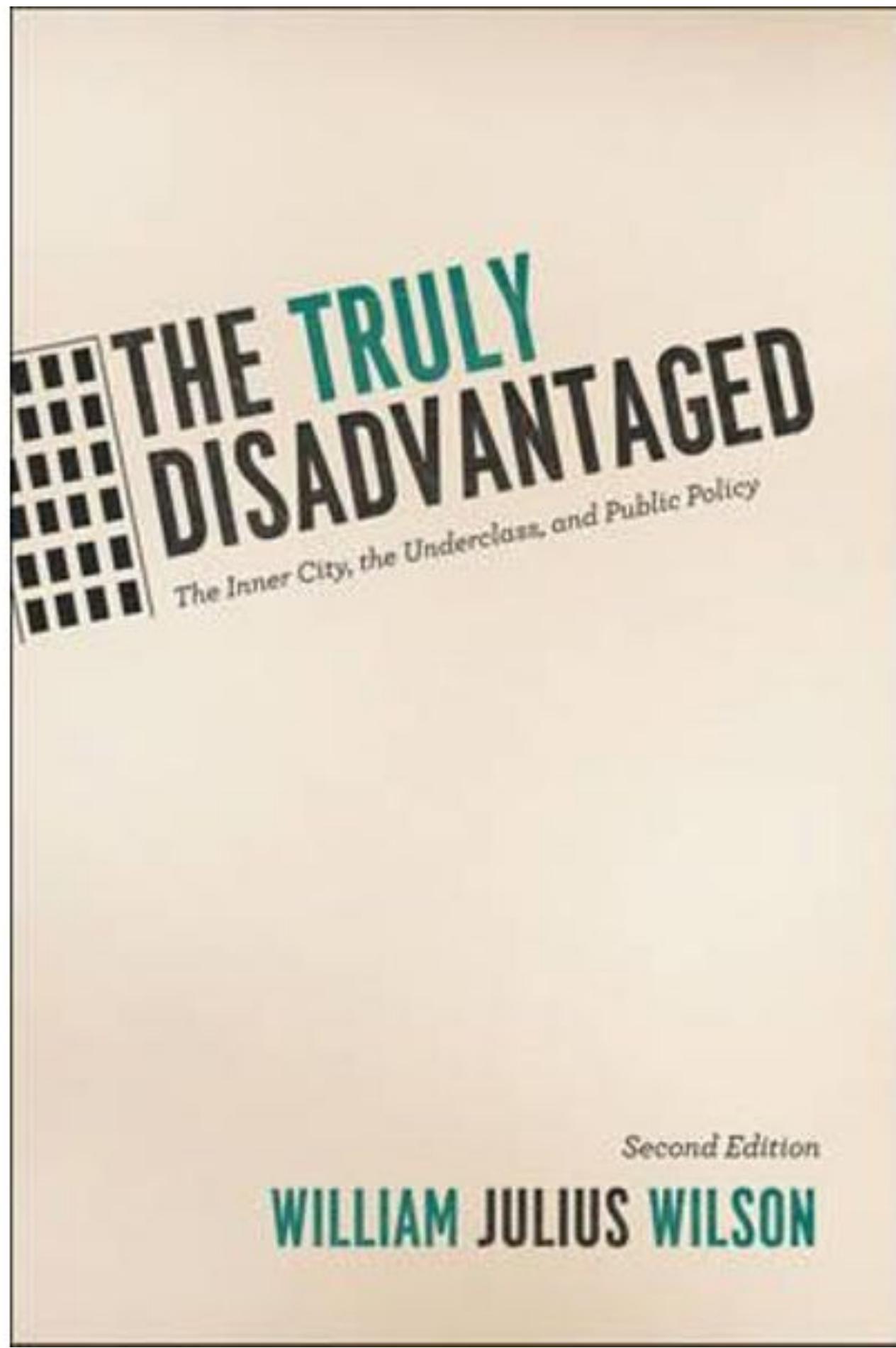


Lecture 11

Looking inside Cities: Spatial Structure and Neighborhoods

11.1 Information, Spatial Selection & the Statistics of Neighborhoods

IUS 6.3



William Julius Wilson, 1987

In "The Truly Disadvantaged," Mr. Wilson takes on conservatives, liberals and civil rights leaders alike as he develops persuasive alternative explanations of what has gone wrong in the inner city and supports them with extensive data and research. For example, he explores the sharpening of class lines among blacks; the proportions of black men with annual incomes over \$25,000 and those with annual incomes under \$5,000 have both increased. Inequality of income is greater now among black families than among white families.

NYtimes review

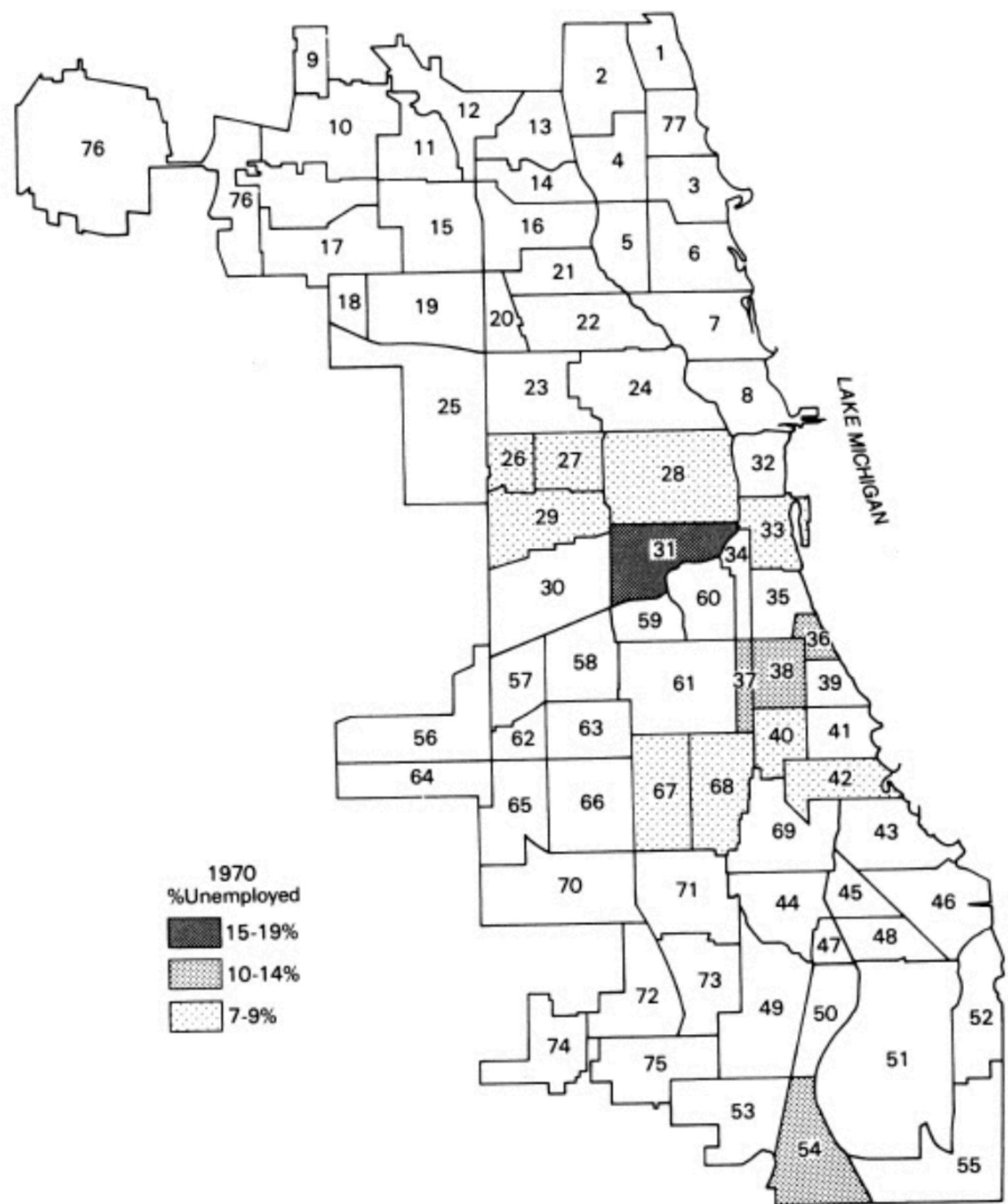


Figure 2.5. Unemployment rates in Chicago Community Areas, 1970.

Source: see [fig. 2.3](#).

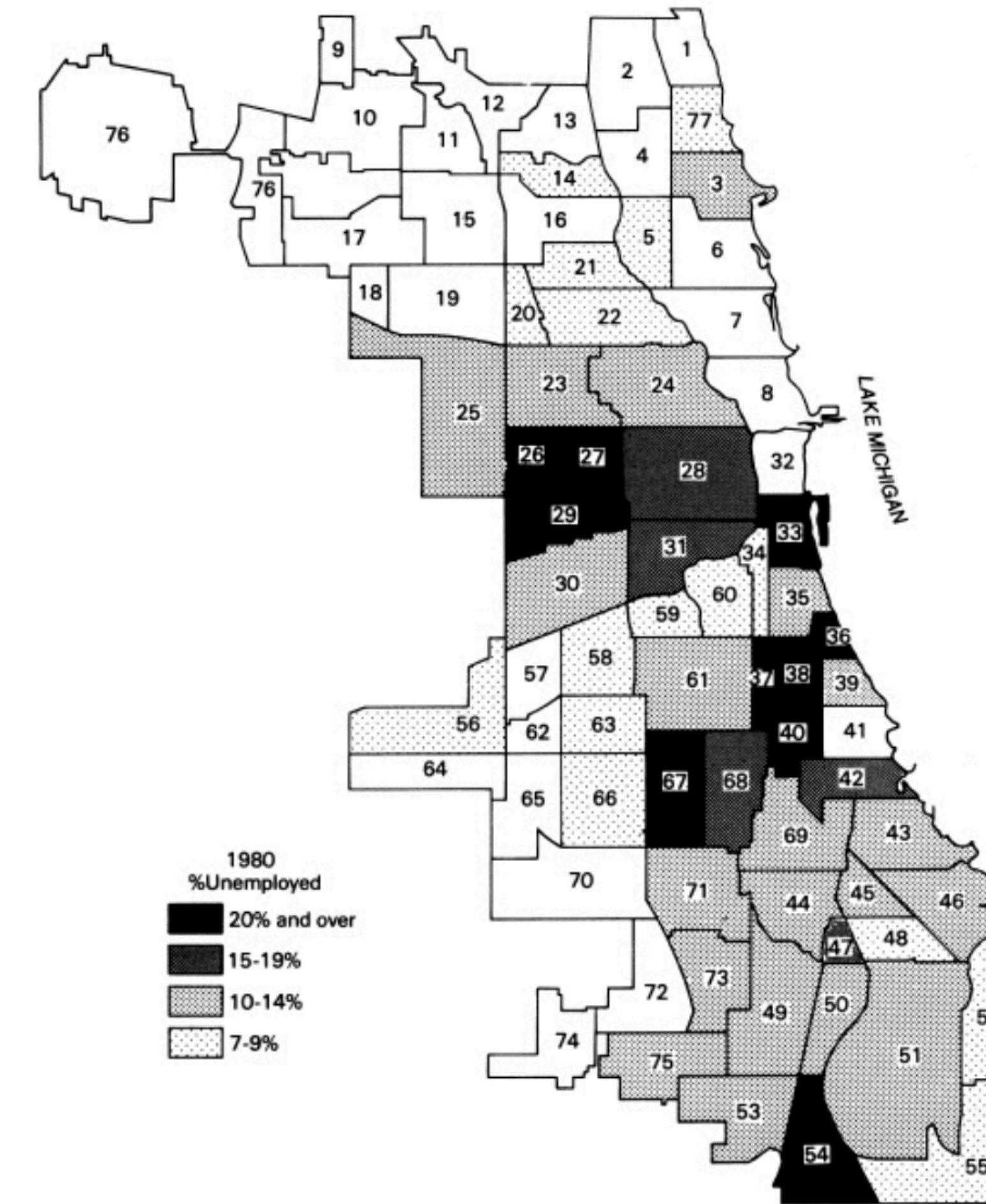
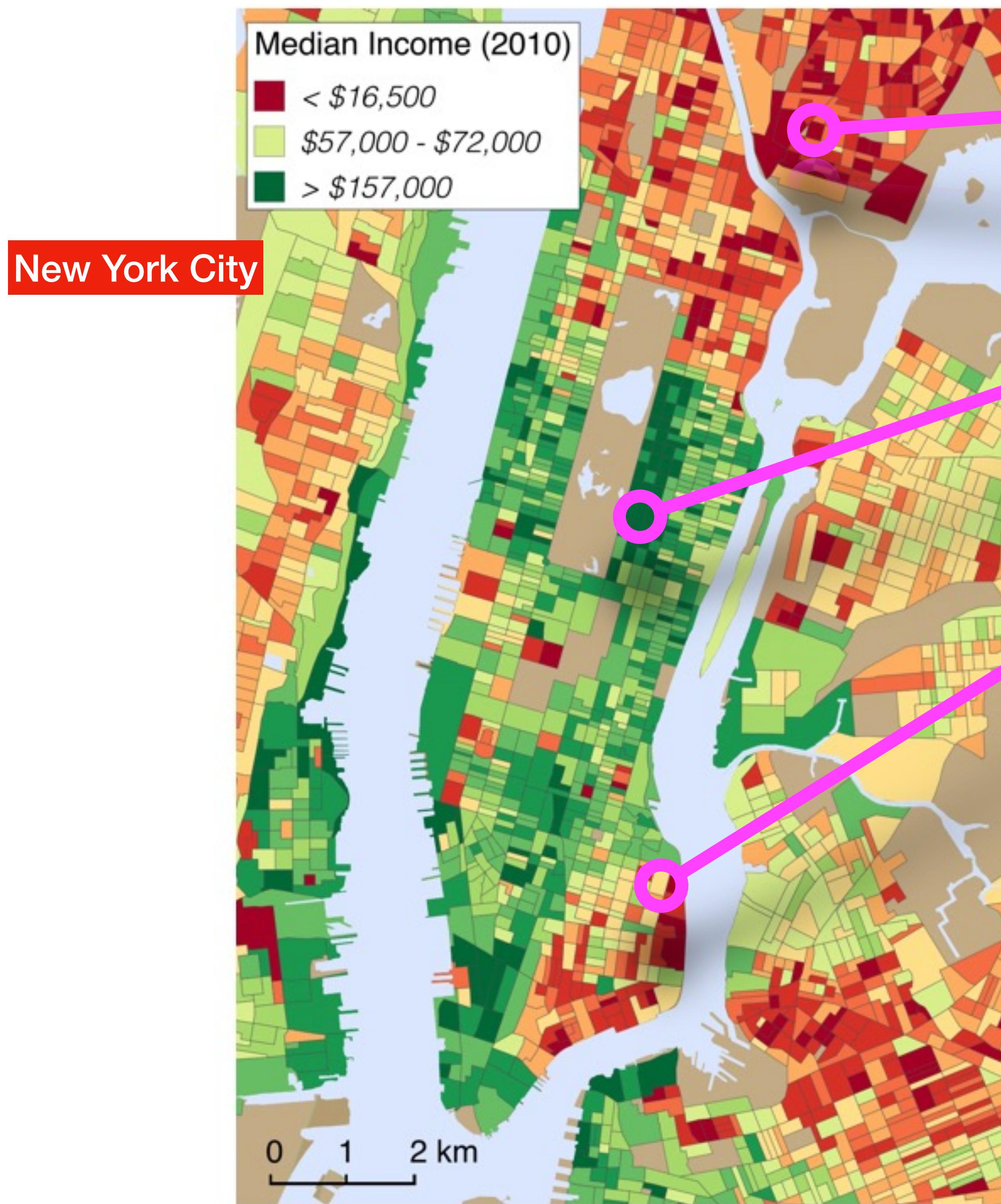
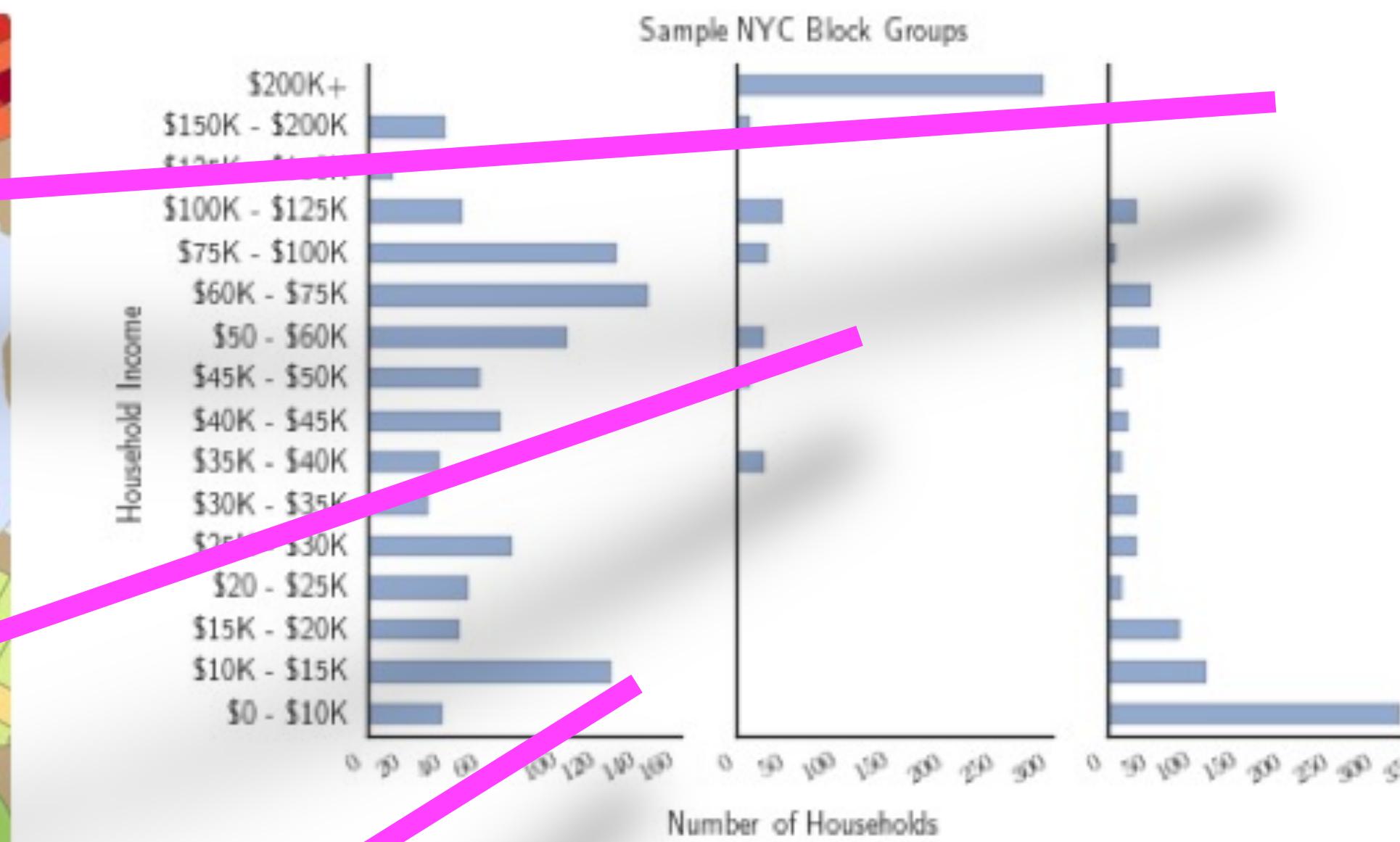
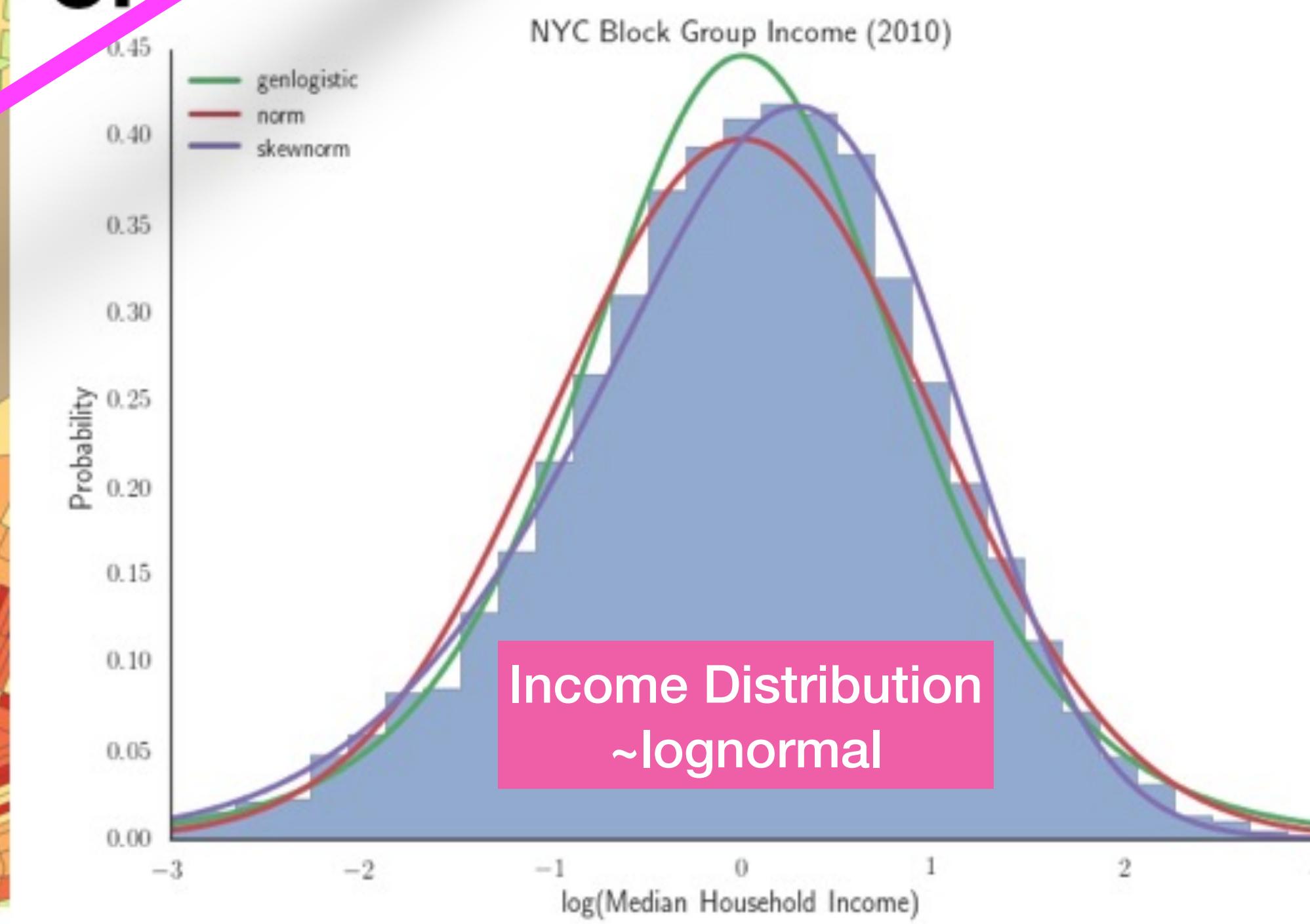


Figure 2.6. Unemployment rates in Chicago Community Areas, 1980.

Source: see [fig. 2.3](#).

A.**B.****C.**

Choice and Information

necessary to get us closer to people and their urban dilemmas

What is “Information” ?

Dictionary

 **in·for·ma·tion**

/ ,ɪnfər'māSH(ə)n/

See definitions in:

All Law Computing Mathematics

noun

1. facts provided or learned about something or someone.
"a vital piece of information"

Similar: details particulars facts figures statistics data knowledge ▾

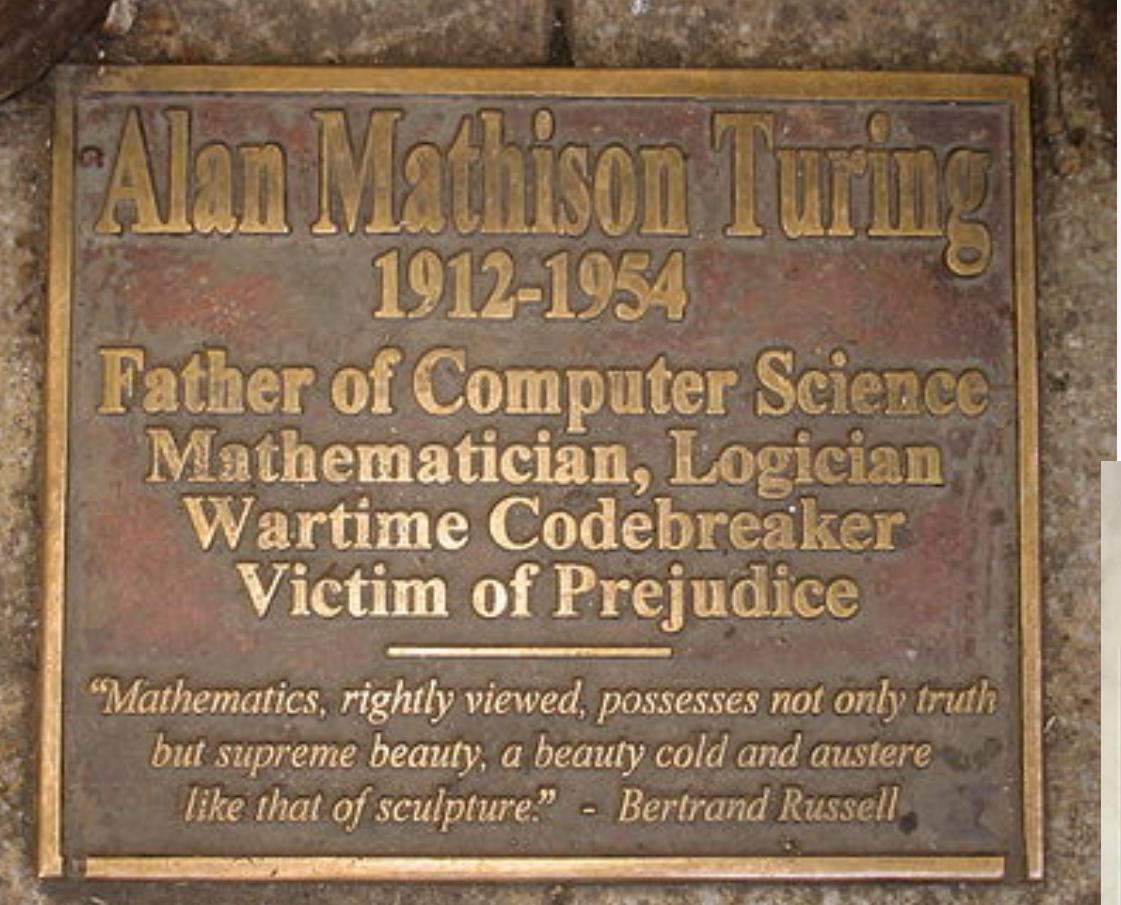
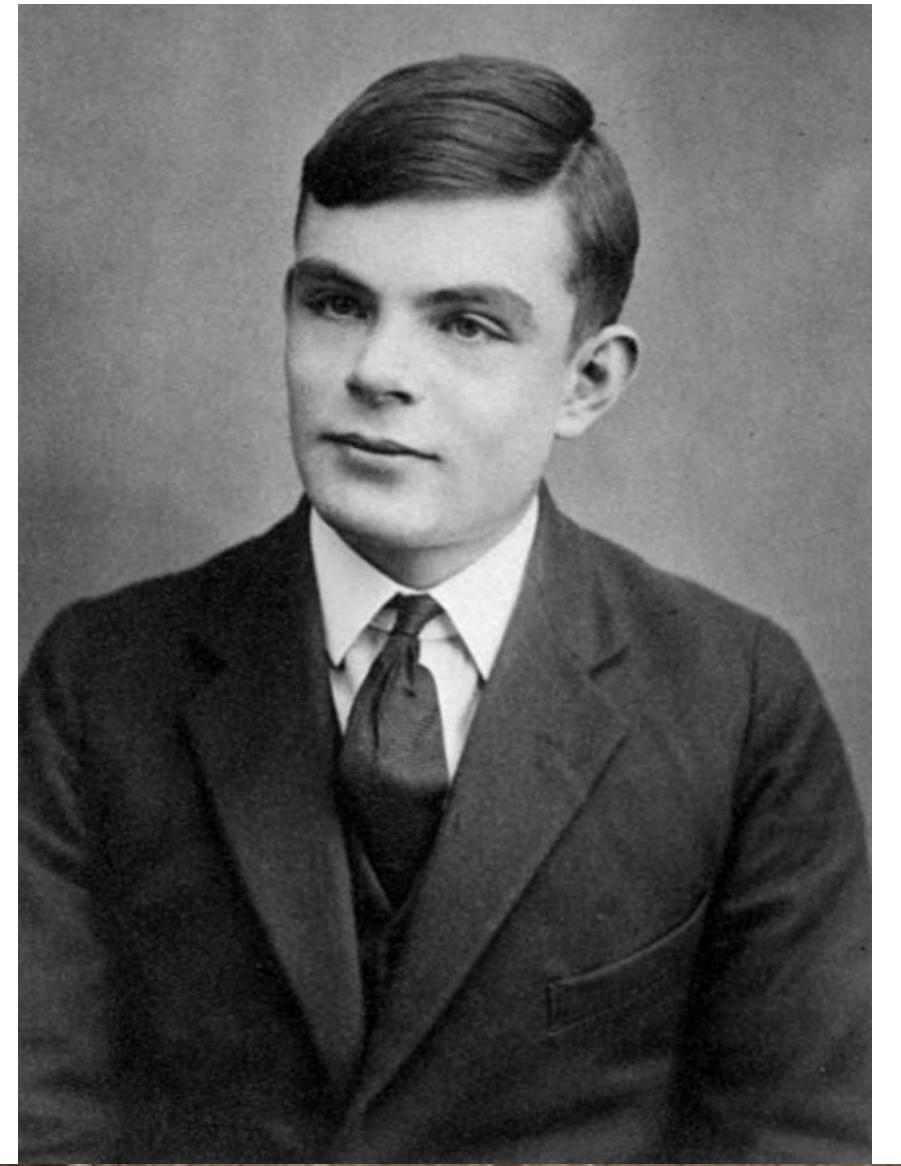
2. what is conveyed or represented by a particular arrangement or sequence of things.
"genetically transmitted information"



**Information is a relative quantity
it is reciprocal:**
you change the city; it changes you



patterns convey information



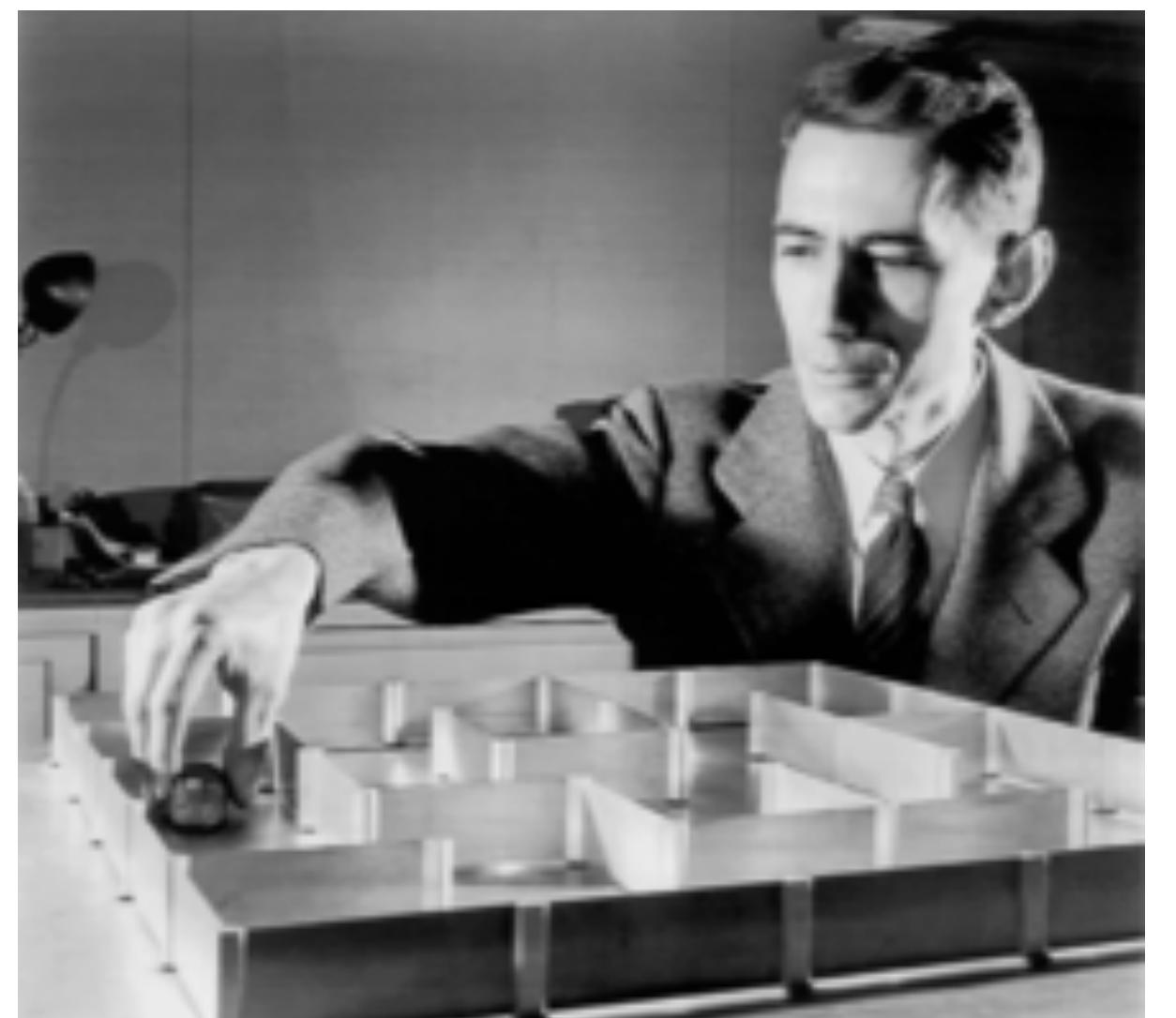
A latent idea

©Crown, courtesy of Director, GCHQ

Cryptography: Bletchley Park



Bell Labs: inventing the telephone



Claude Shannon

The breakthrough paper:

Reprinted with corrections from *The Bell System Technical Journal*,
Vol. 27, pp. 379–423, 623–656, July, October, 1948.

A Mathematical Theory of Communication

By C. E. SHANNON

INTRODUCTION

THE recent development of various methods of modulation such as PCM and PPM which exchange bandwidth for signal-to-noise ratio has intensified the interest in a general theory of communication. A basis for such a theory is contained in the important papers of Nyquist¹ and Hartley² on this subject. In the present paper we will extend the theory to include a number of new factors, in particular the effect of noise in the channel, and the savings possible due to the statistical structure of the original message and due to the nature of the final destination of the information.

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities.

The invention of the “bit”

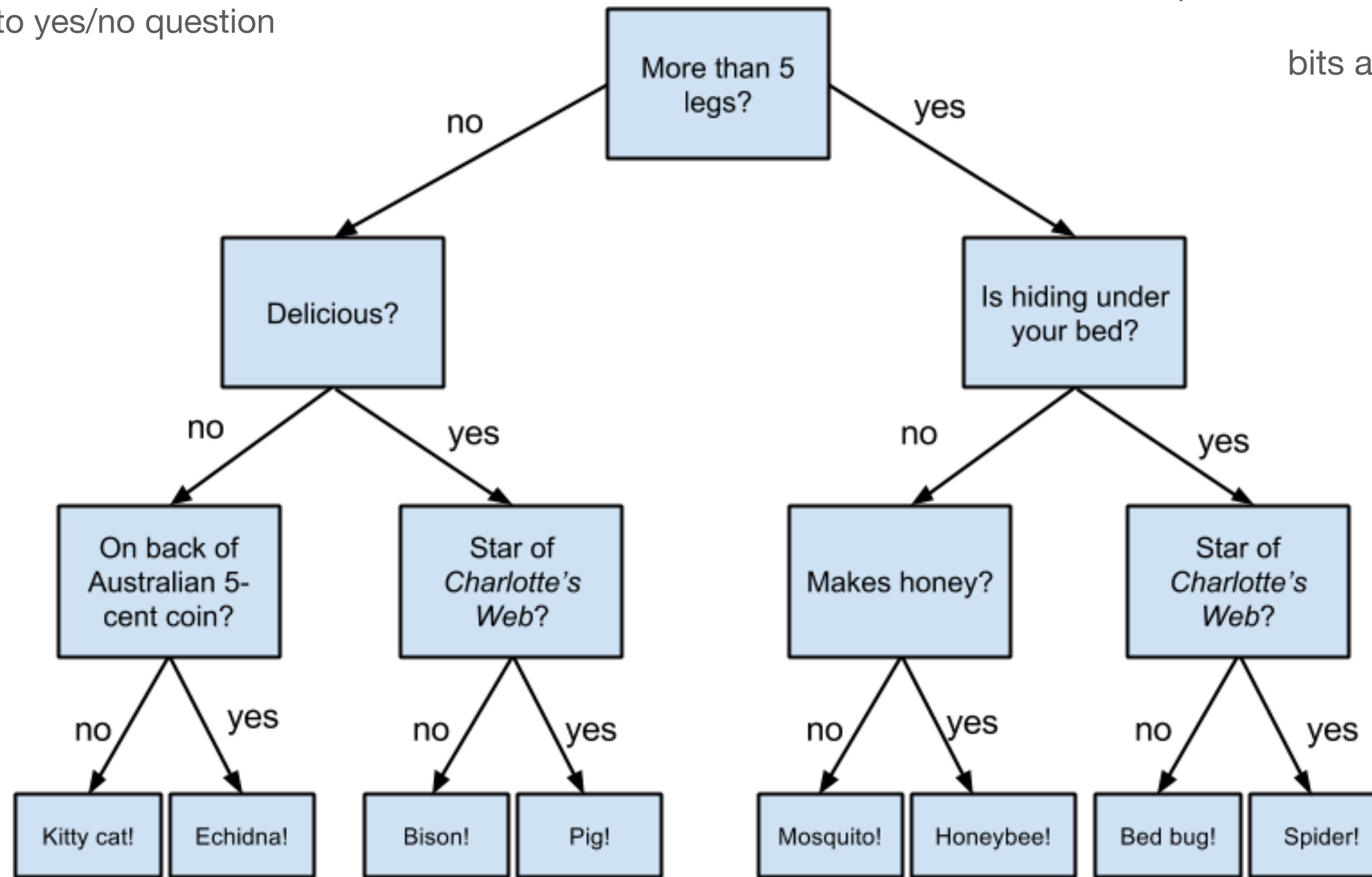
Information and the Game of 20 Questions

what is the “bit”?

“binary unit”: answer to yes/no question

Each question is a decision... = 1bit

bits are units of choice



Information between two variables

$$H(y) = - \sum_i p(y_i) \ln_2 p(y_i)$$

entropy

$$H(y|n) = - \sum_{i,j} p(n_j)p(y_i|n_j) \ln_2 p(y_i|n_j)$$

conditional entropy

$$I(n; y) = H(y) - H(y|n) = \sum_{i,j} p(y_i, n_j) \ln_2 \frac{p(y_i, n_j)}{p(y_i)p(n_j)}$$

mutual information

example:

(fair coin)

$$y = 0101001101 \quad \begin{matrix} \uparrow \\ p(y=0) = p(y=1) = \frac{1}{2} \end{matrix}$$

$$H(y) = 2 \frac{1}{2} \ln_2 2 = 1 \text{ bit}$$

A tip from a friend

$x = 0101001101 \quad \text{perfect prediction!}$

$$\begin{aligned} p(y=1|x=1) &= p(y=0|x=0) = 1 & H(y|x) &= 0 \\ p(x=1) &= p(x=0) = \frac{1}{2} & I(y;x) &= H(y) - H(y|x) = H(y) = 1 \end{aligned}$$

Information between two variables

$$H(y) = - \sum_i p(y_i) \ln_2 p(y_i)$$

$$H(y \mid n) = - \sum_{i,j} p(n_j) p(y_i \mid n_j) \ln_2 p(y_i \mid n_j)$$

$$I(n; y) = H(y) - H(y \mid n) = \sum_{i,j} p(y_i, n_j) \ln_2 \frac{p(y_i, n_j)}{p(y_i)p(n_j)}$$

example:

(fair coin)

A tip from a friend

$$y = \texttt{0101001101} \quad \begin{matrix} \uparrow \\ 0 \end{matrix} \quad p(y=0) = p(y=1) = \frac{1}{2}$$

$$H(y) = 2 \frac{1}{2} \ln_2 2 = 1 \text{ bit}$$

$$x = \texttt{ababaaabbaba}$$

perfect prediction!

$$p(y=1 \mid x=b) = p(y=0 \mid x=a) = 1$$

$$H(y \mid x) = 0$$

$$p(x=a) = p(x=b) = \frac{1}{2}$$

$$I(y; x) = H(y) - H(y \mid x) = H(y) = 1$$

Information between two variables

example:
(fair coin)

$$y = \begin{array}{cccccccccc} 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{array}$$

$$p(y = 0) = p(y = 1) = \frac{1}{2}$$

$$H(y) = 2 \frac{1}{2} \ln_2 2 = 1 \text{ bit}$$

A **noisy tip from a friend**

$$x = \begin{array}{cccccccccc} a & a & a & b & b & a & b & b & a & b & a \end{array}$$

imperfect prediction!

$$p(y = 1 \mid x = b) = p(y = 0 \mid x = a) = \frac{4}{5}$$

$$p(y = 1 \mid x = a) = p(y = 0 \mid x = b) = \frac{1}{5}$$

$$p(x = a) = p(x = b) = \frac{1}{2}$$

$$H(y \mid x) = 0.26$$

$$I(y; x) = H(y) - H(y \mid x) = H(y) = 1 - 0.26 = 0.74 < H(y)$$

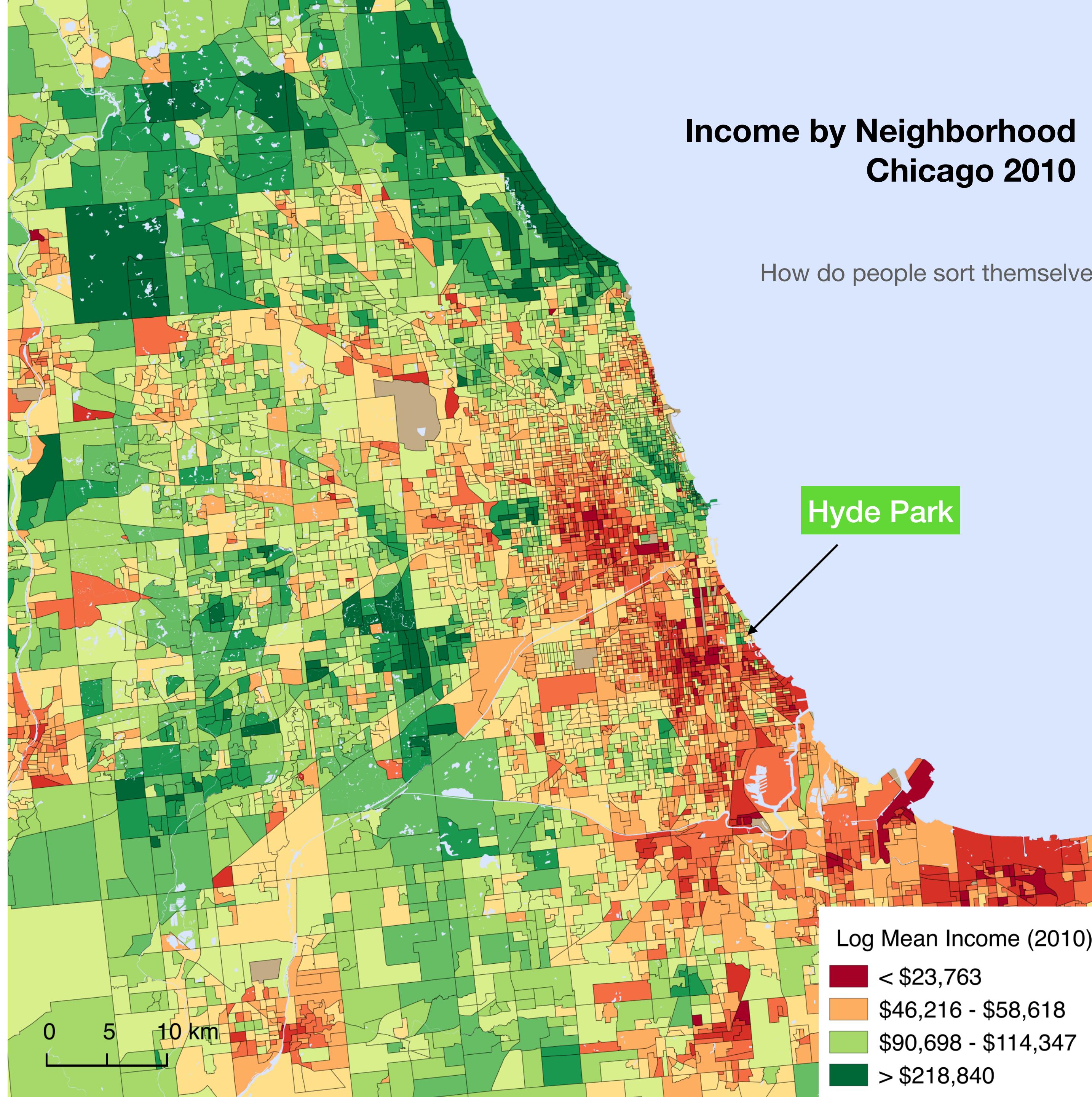
The tip contains predictive information, but not enough for *perfect* prediction!

This is true of most situations in real life!

Income by Neighborhood Chicago 2010

How do people sort themselves/ are sorted?

Hyde Park

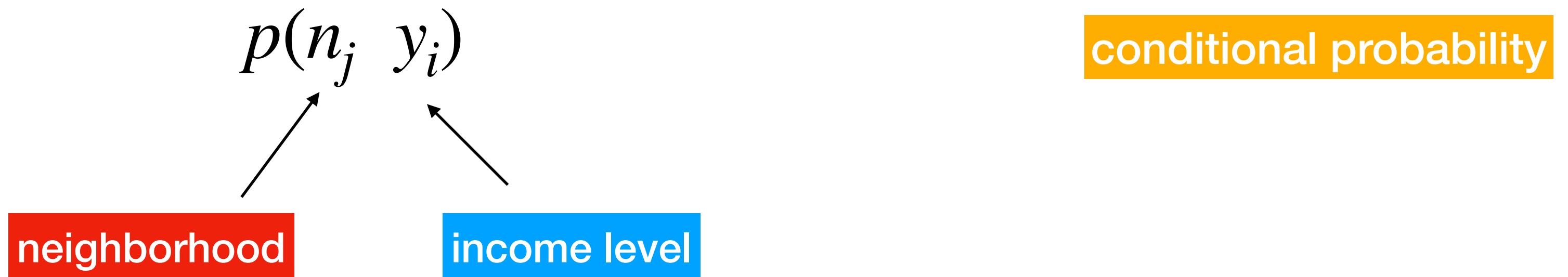


Back to cities: Pattern

If I know where you live, how much do I know about your income?



If I know your income, can I predict where you live?



Bayes Theorem

Remember that probabilities combine as :

$$p(A, B) = p(A | B)p(B)$$

$$p(A | B) = \frac{p(B | A)}{p(B)} p(A)$$

optimal way of learning !!

All theories of learning are Bayesian: Evolution, Cognition, Child Development, Statistical Estimation, Machine Learning, AI

For neighborhoods:

probability of income level given neighborhood

$$p(y_i | n_j) = \frac{p(n_j | y_i)}{p(n_j)} p(y_i) \quad \text{probability of income level}$$

$$w_{ij} = \frac{p(n_j | y_i)}{p(n_j)} = \frac{p(y_i | n_j)}{p(y_i)} = \frac{p(n_j, y_i)}{p(y_i)p(n_j)}$$

strength of neighborhood selection

Information between two variables (Neighborhood vs. income level)

$$I(n; y) = H(y) - H(y|n) = \sum_{i,j} p(y_i, n_j) \ln_2 \frac{p(y_i, n_j)}{p(y_i)p(n_j)}$$

This is called the “mutual information”

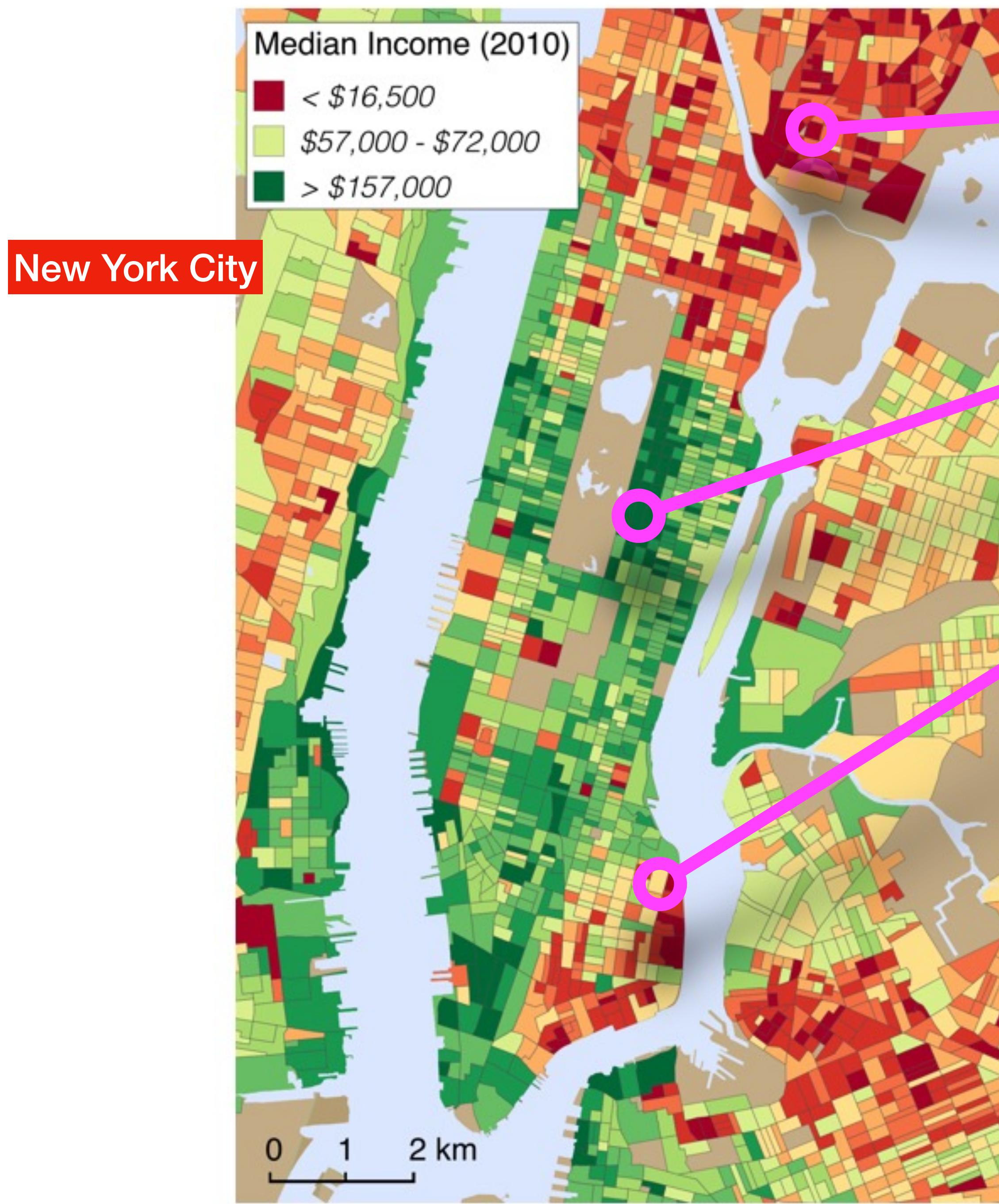
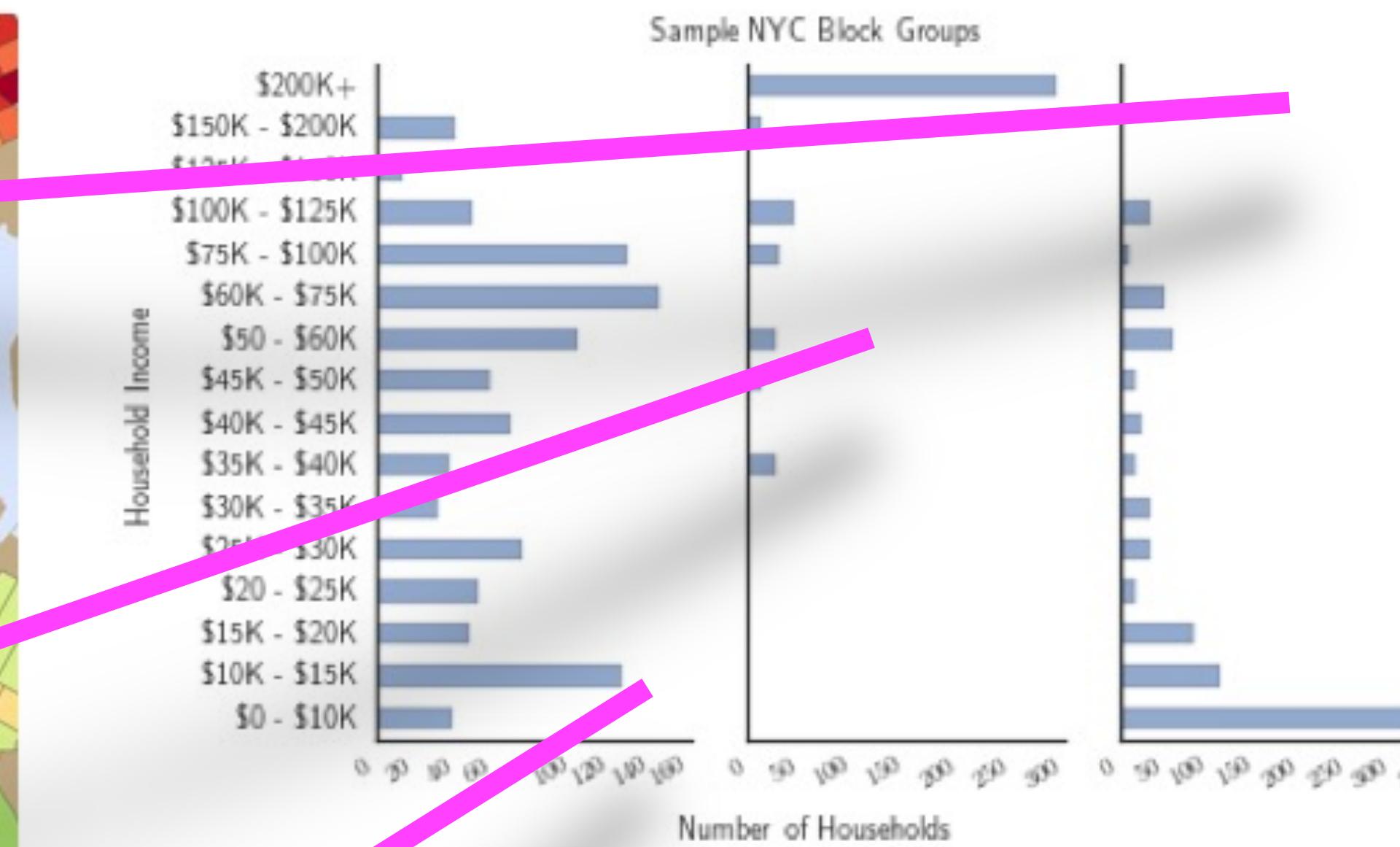
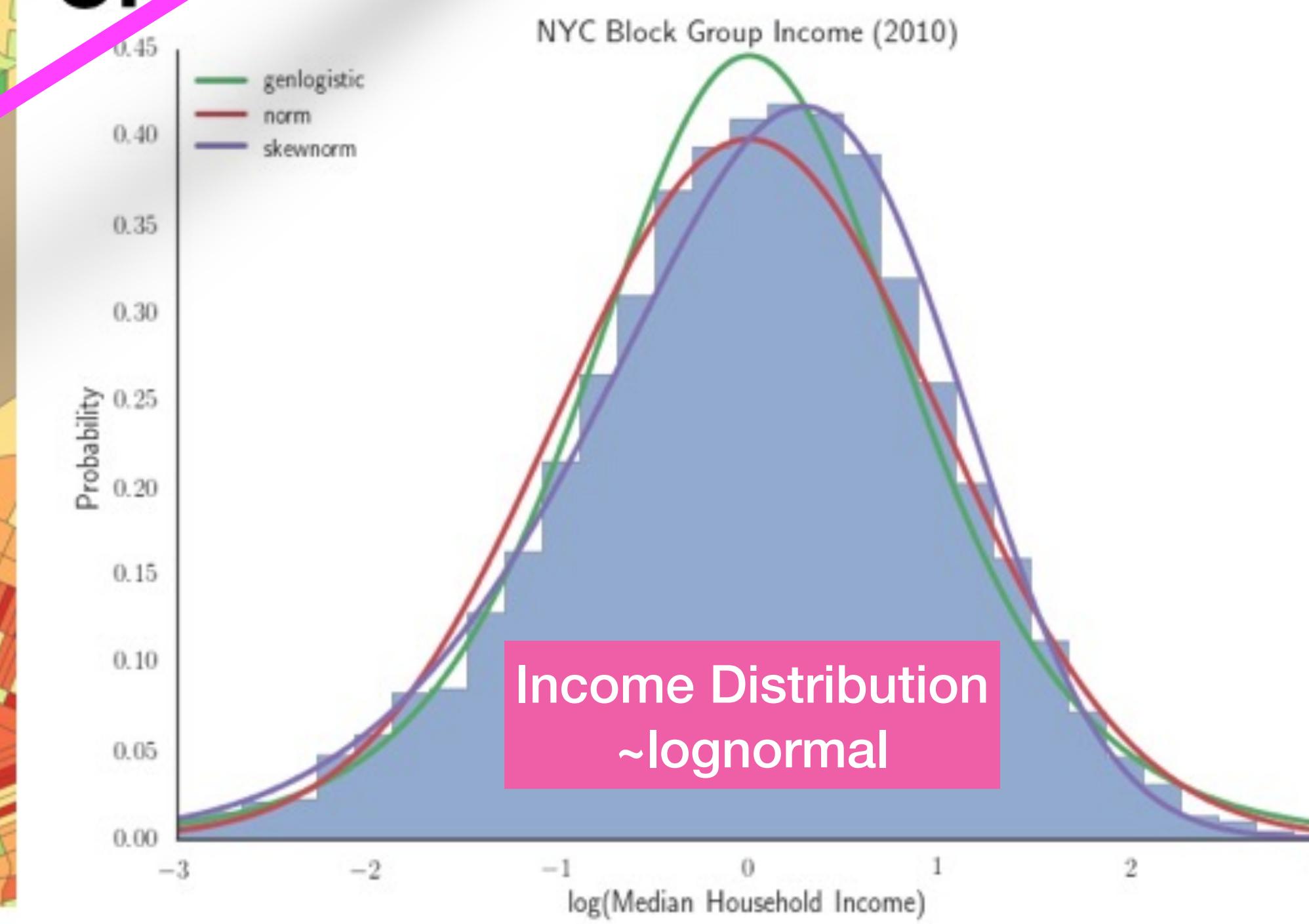
The information in each specific neighborhood:

$$D(n_j) = \sum_i p(y_i|n_j) \ln_2 \frac{p(y_i|n_j)}{p(y_i)}$$

This is called the “Kullback-Leibler (KL) divergence”

The information in each specific income group:

$$D(y_i) = \sum_j p(n_j|y_i) \ln_2 \frac{p(n_j|y_i)}{p(n_j)}$$

A.**B.****C.**

Information & Spatial Selection

$$p(y_\ell | n_j) = w_{\ell,j} p(y_\ell)$$

Income distribution for neighborhood i

Income distribution for whole city population

$$p(y_\ell | n_j) = \frac{p(n_j | y_\ell) p(y_\ell)}{p(n_j)} \rightarrow w_{\ell,j} = \frac{p(n_j | y_\ell)}{p(n_j)} = \frac{p(y_\ell | n_j)}{p(y_\ell)} = \frac{p(n_j, y_\ell)}{p(y_\ell) p(n_j)}$$

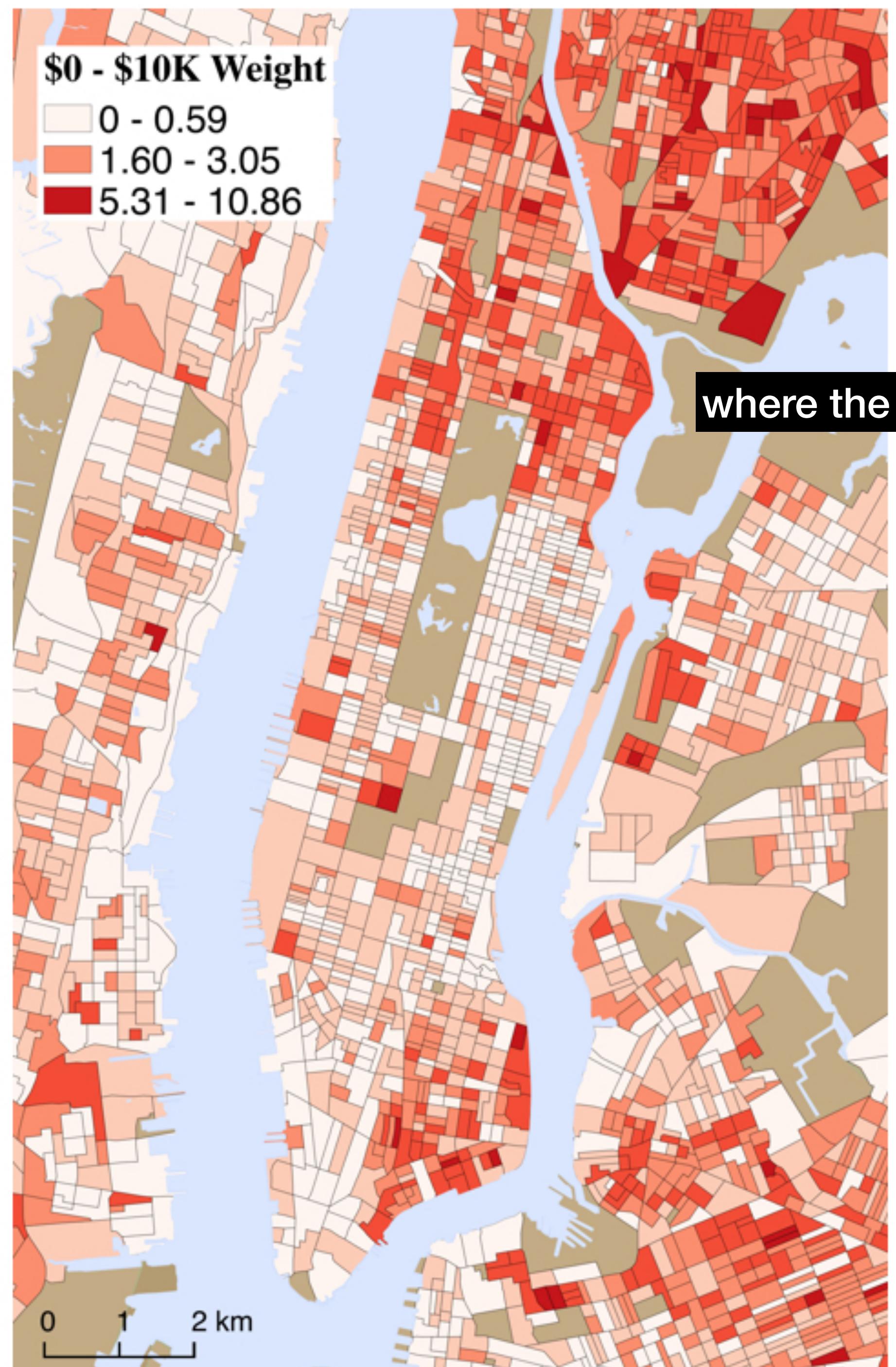
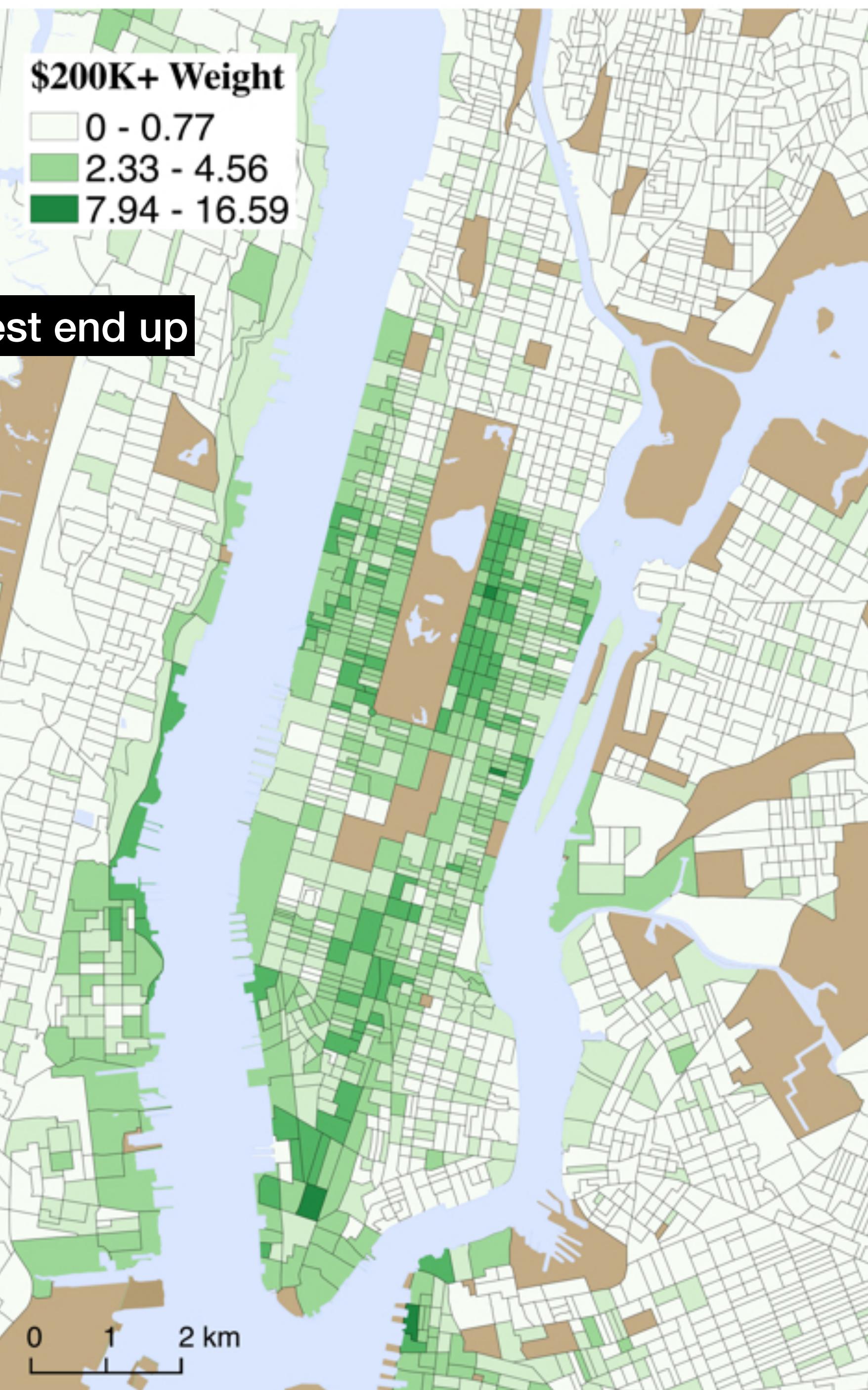
Bayes Theorem

$$\langle \ln w_j \rangle = \sum_\ell p(y_\ell | n_j) \ln \frac{p(y_\ell | n_j)}{p(y_\ell)} = D_{\text{KL}} [p(y_\ell | n_j) || p(y_\ell)]$$

Information
to explain neighborhood j

$$\langle \ln w \rangle = \sum_{i,l} p(n_i, y_\ell) \ln w_{i,l} = I(n; y).$$

Information
between neighborhoods and income
for a city



For 15 years, David Koch lived at the world's "richest building"

Brokerage sources speculate that Koch's unit may list for sale after construction is complete at his recently purchased UES townhouse

TRD New York / By [E.B. Solomont](#)

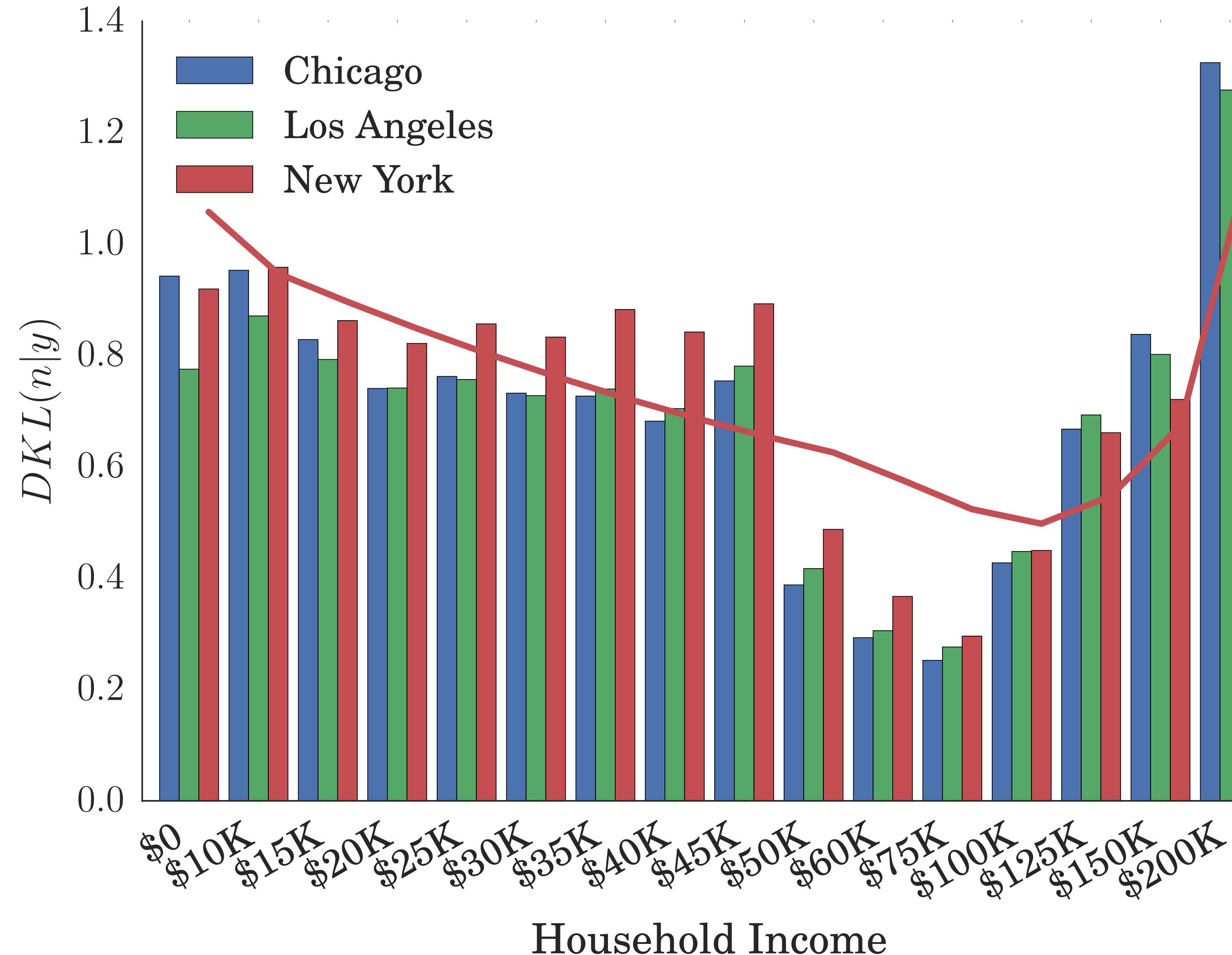
August 26, 2019 07:00 AM



Clockwise from left: John D. Rockefeller, Izzy Englander, Steven Mnuchin, David Koch, Jacqueline Bouvier, and William Zeckendorf (Credit: Getty Images and StreetEasy)

Before paying \$40 million for an Upper East Side townhouse last year, David Koch was among a handful of billionaires living at 740 Park Avenue.

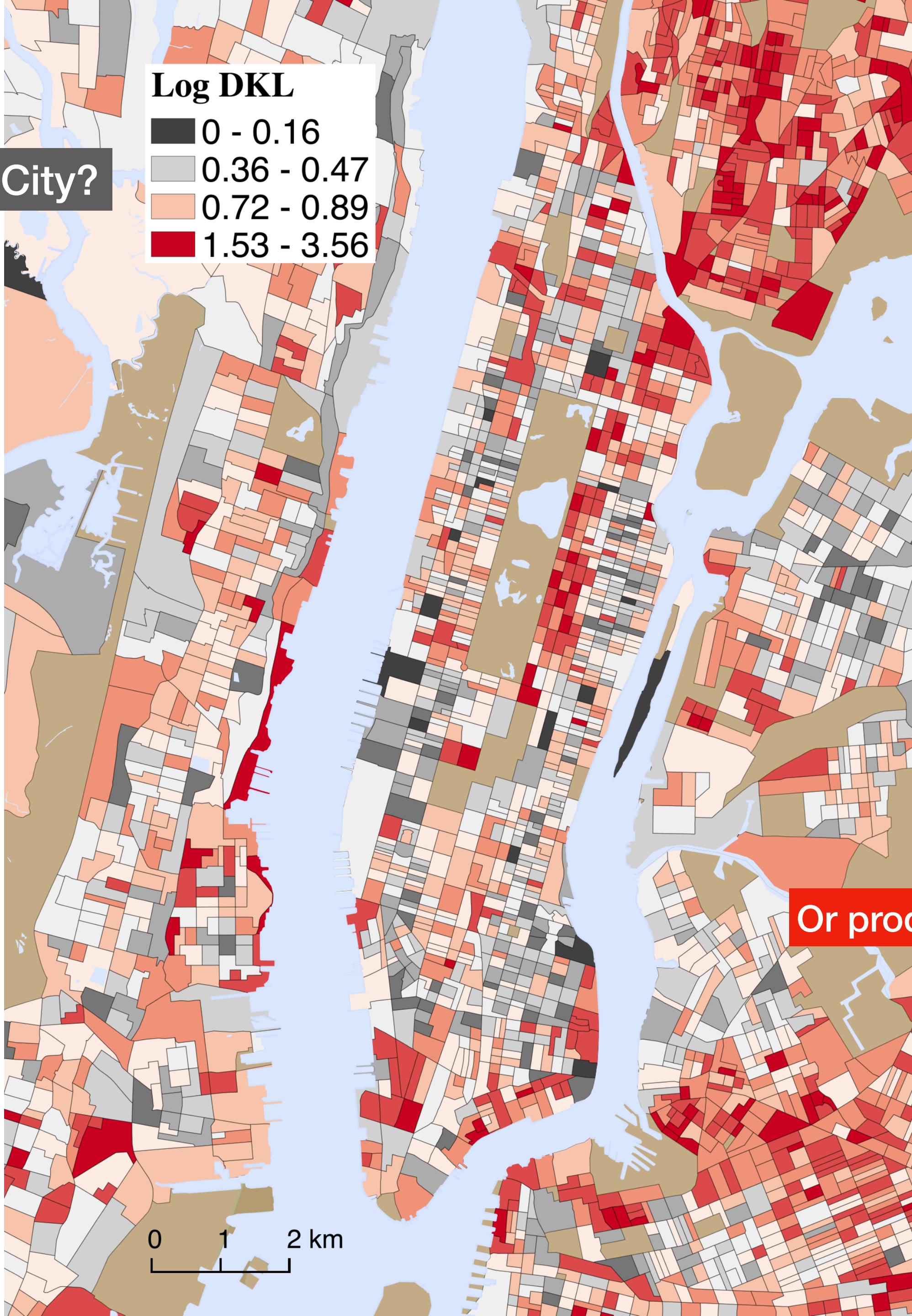
which income group is most segregated spatially?



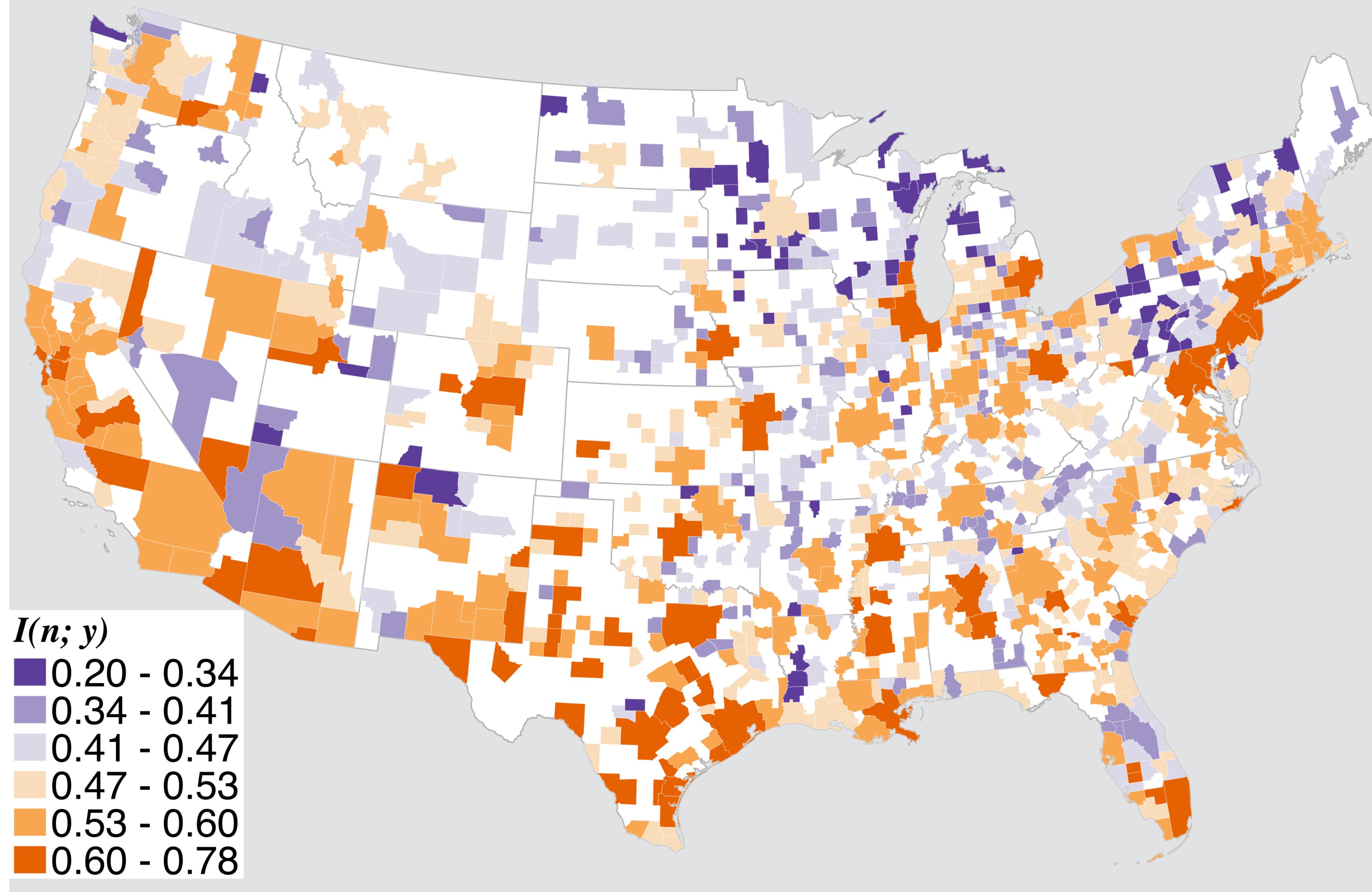
Microcosms of the City?

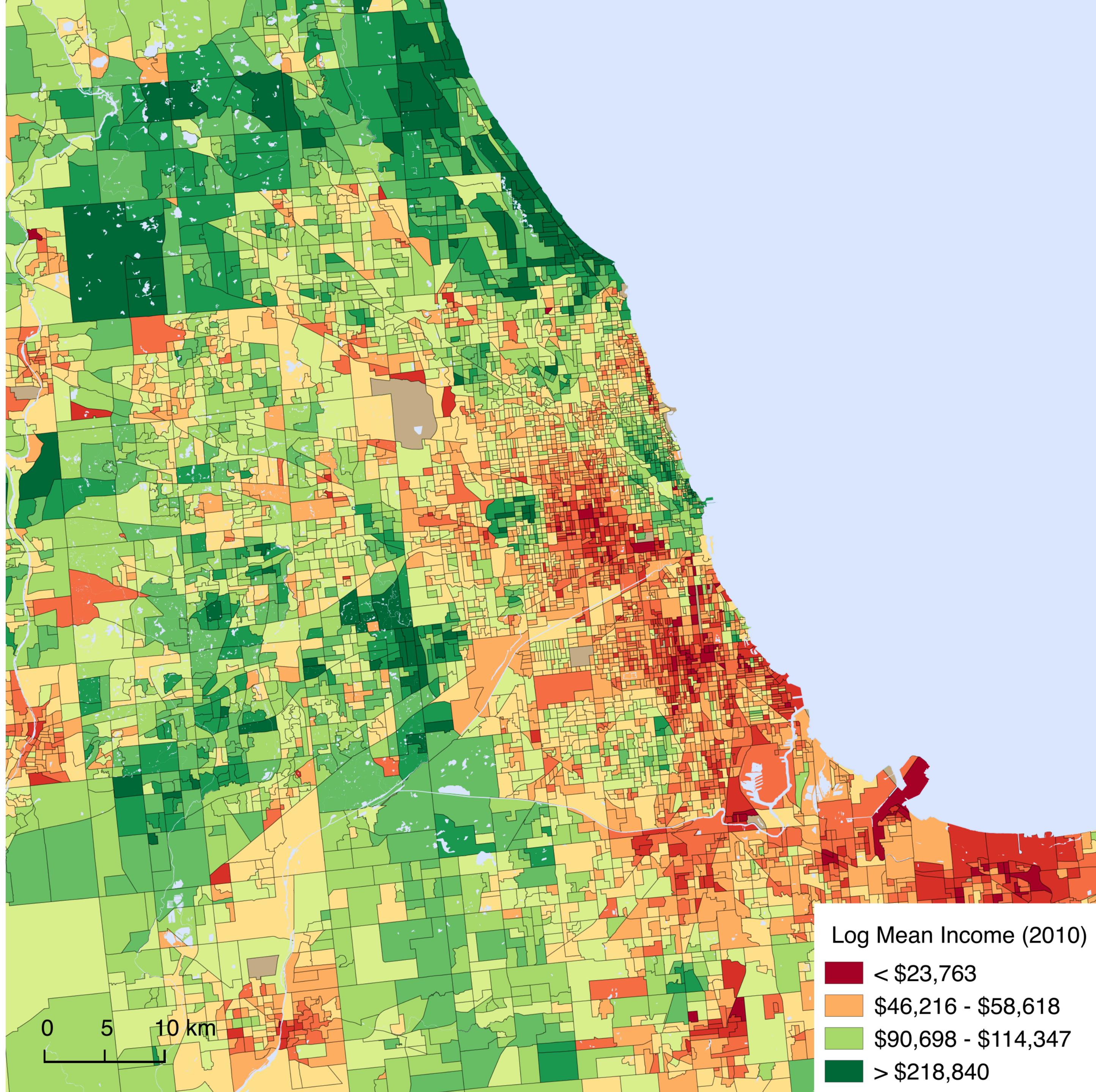
Log DKL

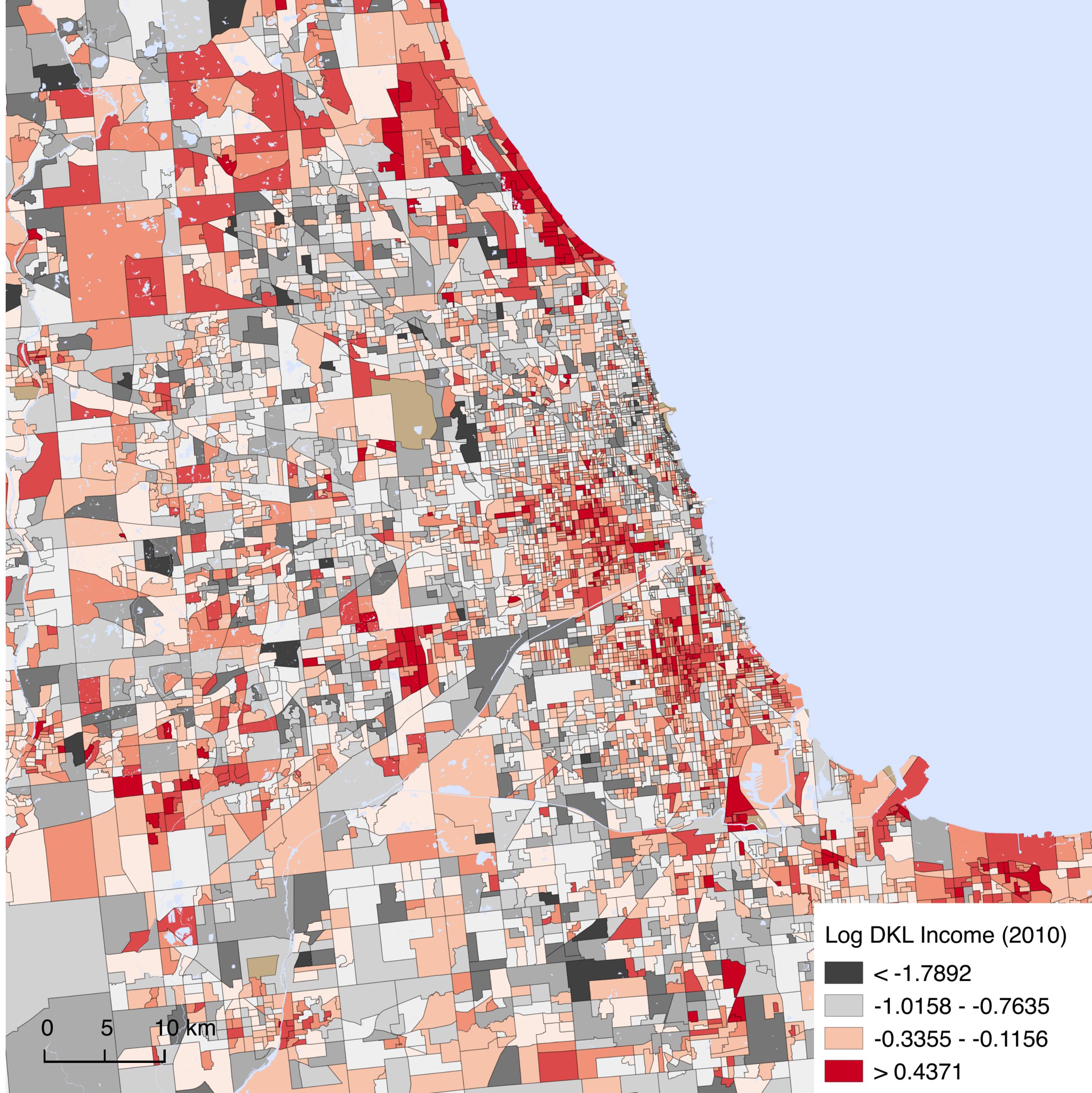
- 0 - 0.16
- 0.36 - 0.47
- 0.72 - 0.89
- 1.53 - 3.56



In which cities do neighborhoods tell me more about income ?







General features of household neighborhood sorting by income in US Metros

-There are many types of neighborhoods:

Segregated, Integrated, Rich and Poor,...

-The greater the city-wide inequality, the more polarized neighborhoods tend to be

-Amounts of segregation and neighborhood effects vary considerably

from place to place and city to city