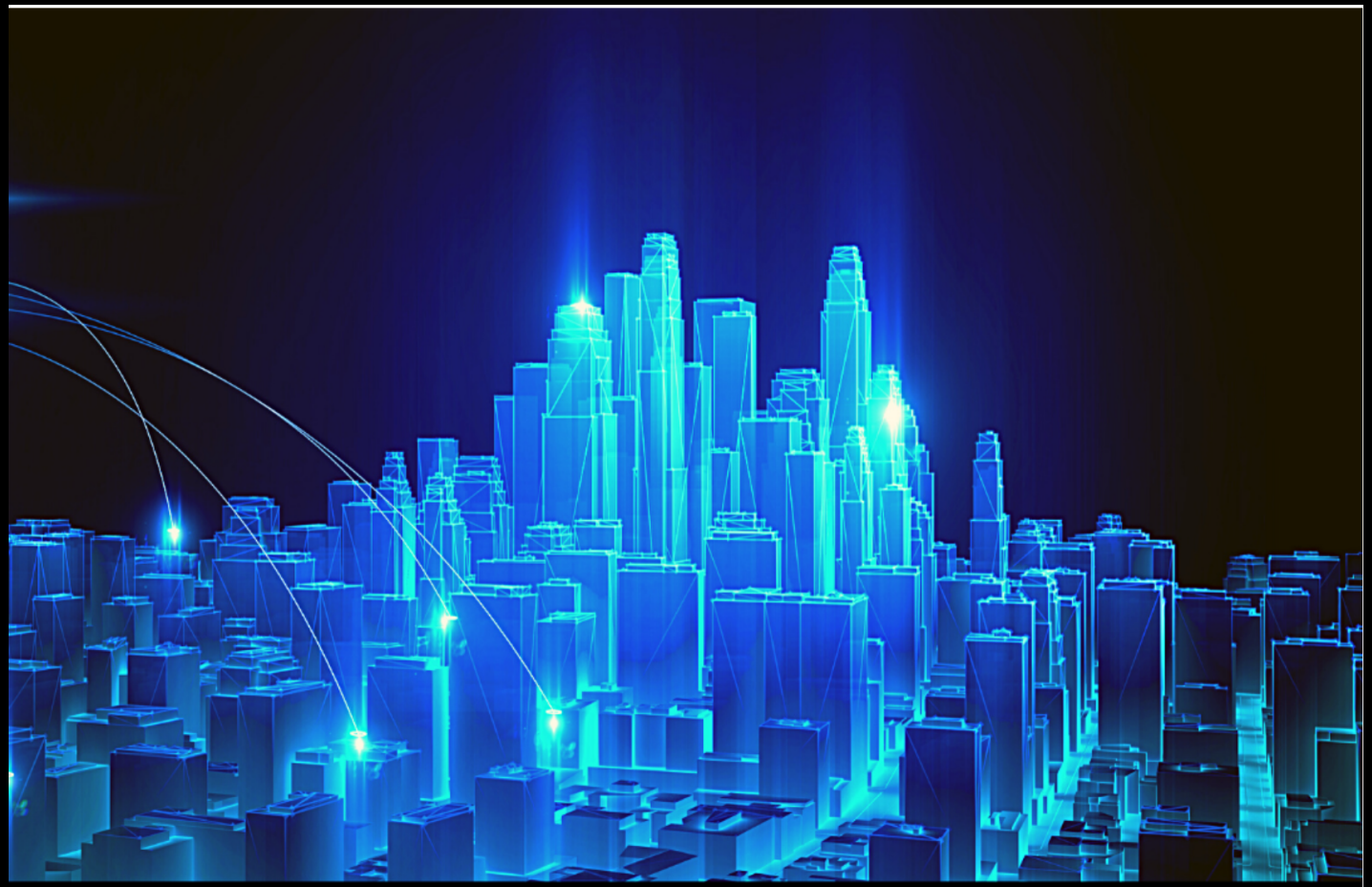


CITIES AND DATA

WEDNESDAY, AUGUST 19TH 6:00-7:00PM

CITY PLANNING & URBAN AFFAIRS
METROPOLITAN COLLEGE
BOSTON UNIVERSITY

$$I_i = \frac{x_i - \bar{X}}{S_i^2} \sum_{j=1, j \neq i}^{n} w_{i,j}(x_j - \bar{X})$$

$$S_i^2 = \frac{\sum_{j=1, j \neq i}^{n} (x_j - \bar{X})^2}{n - 1}$$

1. Is the fascination with data for urban planning new?

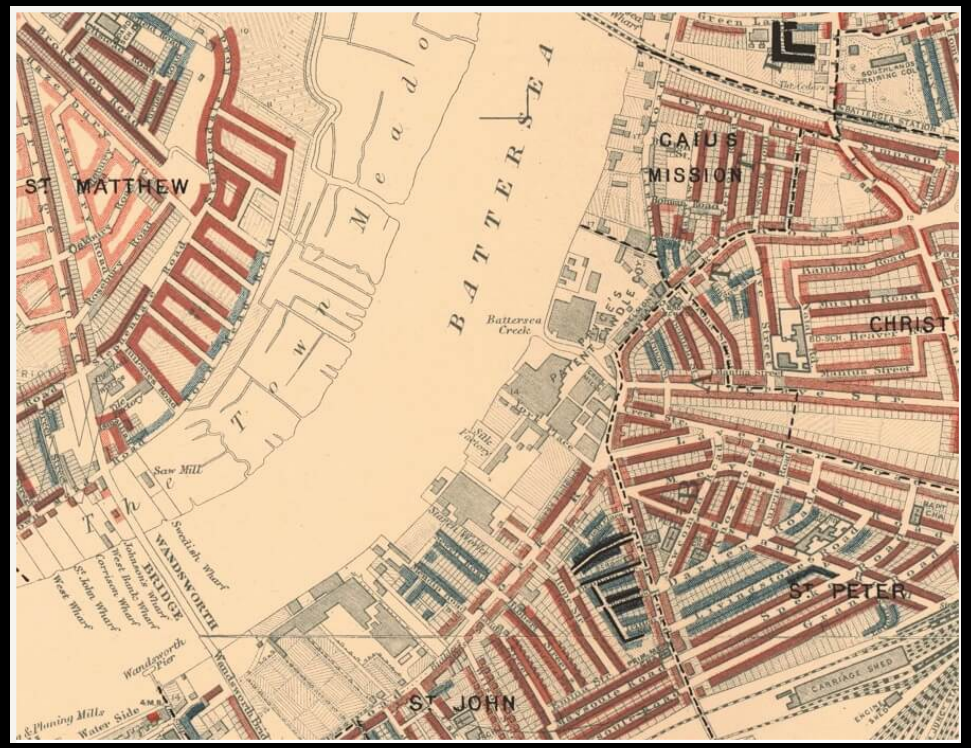
2. How much and what kinds of data out there today?

3. Why is context as important as the data itself?

4. What do we do with data?

1. The fascination with data for urban planning new?

Is the fascination with data for urban planning new? No

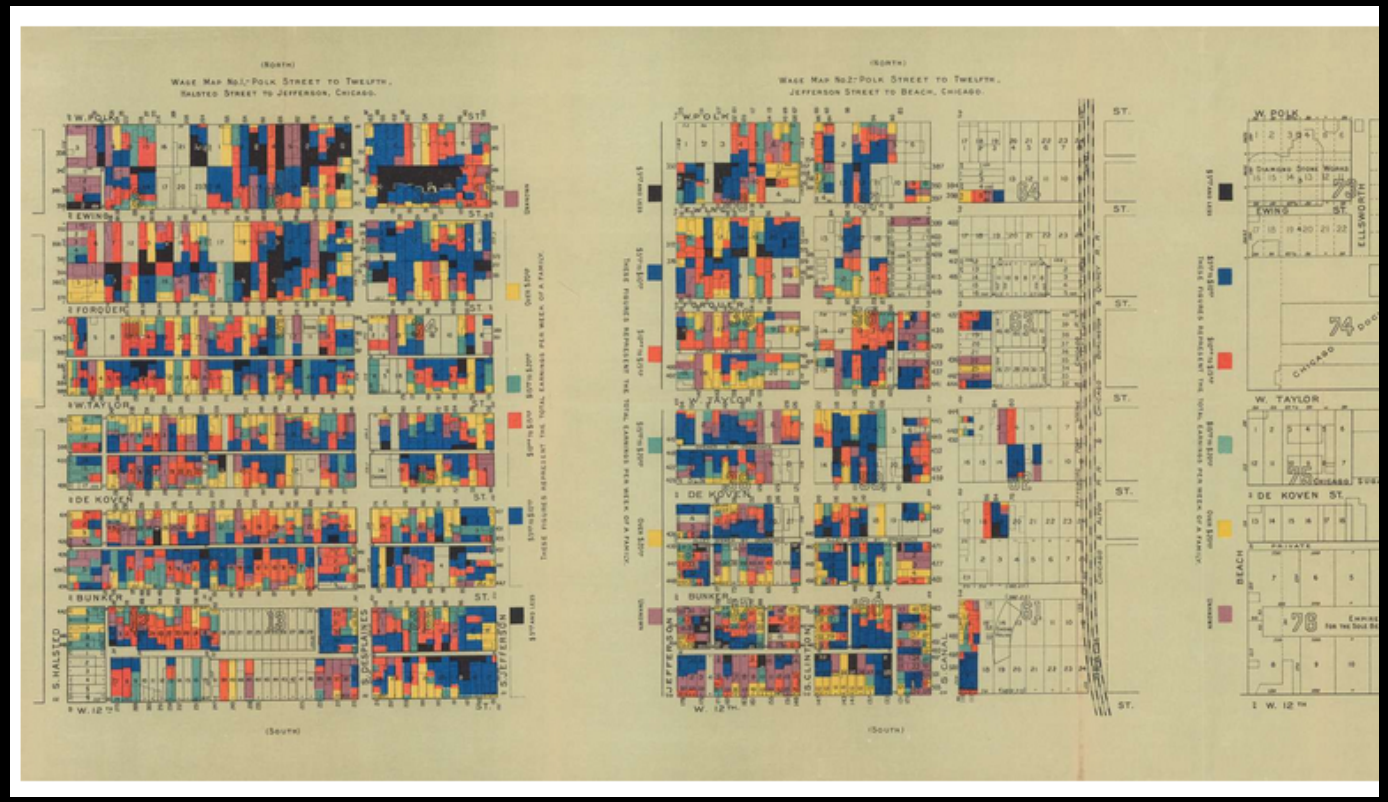


Charles Booth's Inquiry into Life and Labour in London (1886-1903)

Classification	Colour	
Lowest class. Vicious, semi-criminal.	Black	
Very poor, casual. Chronic want.	Dark blue	
Poor. 18s. to 21s. a week for a moderate family.	Light blue	
Mixed. Some comfortable others poor.	Purple	
Fairly comfortable. Good ordinary earnings.	Pink	
Middle class. Well-to-do.	Red	
Upper-middle and upper classes. Wealthy.	Yellow	

https://booth.lse.ac.uk/learn-more/what-were-the-poverty-maps

Is the fascination with data for urban planning new?

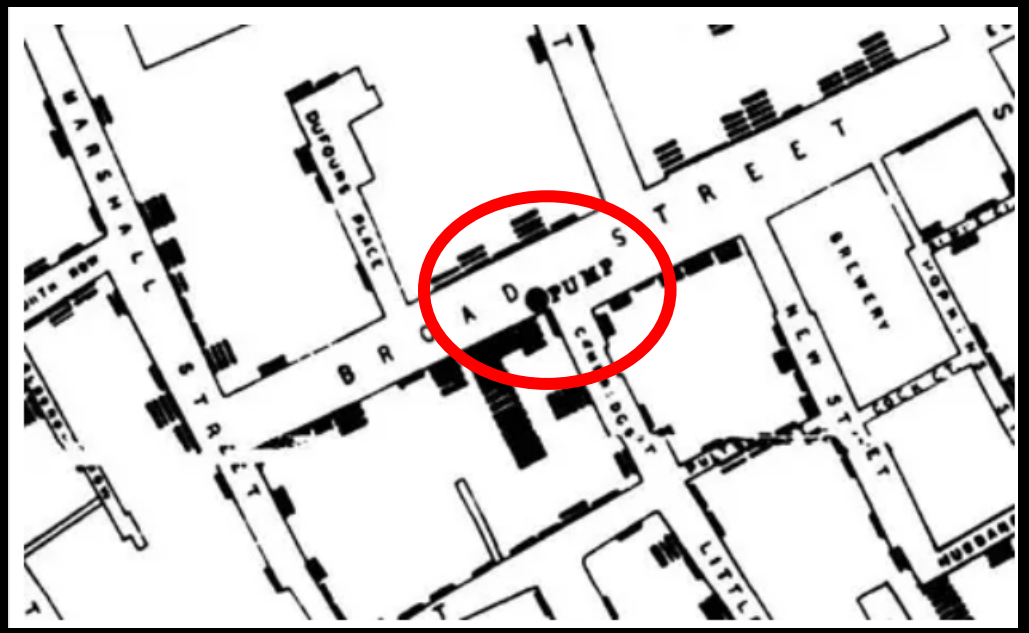


Florence Kelley -- Chicago Wage Maps (1895)

https://atlasofplaces.com/cartography/chicago-wage-maps/

Is the fascination with data for urban planning new?

John Snow's data about cholera deaths and water pumps

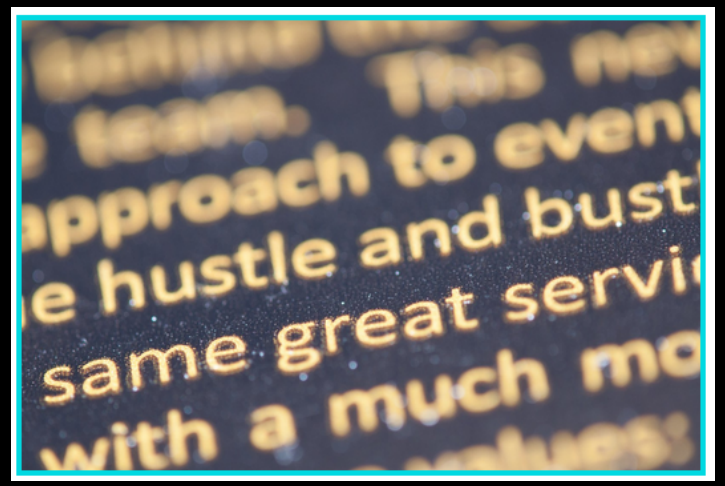


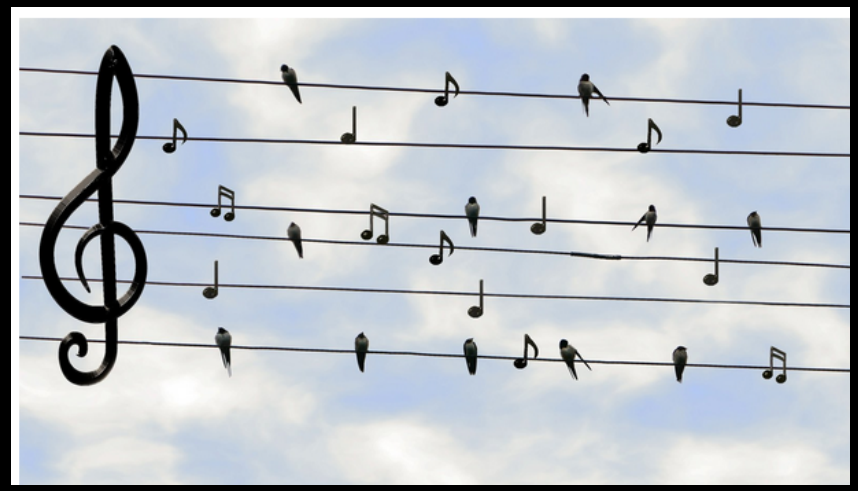
https://www.rcseng.ac.uk/library-and-publications/library/blog/mapping-disease-john-snow-and-cholera/

2. How much and what kind of data out there today?

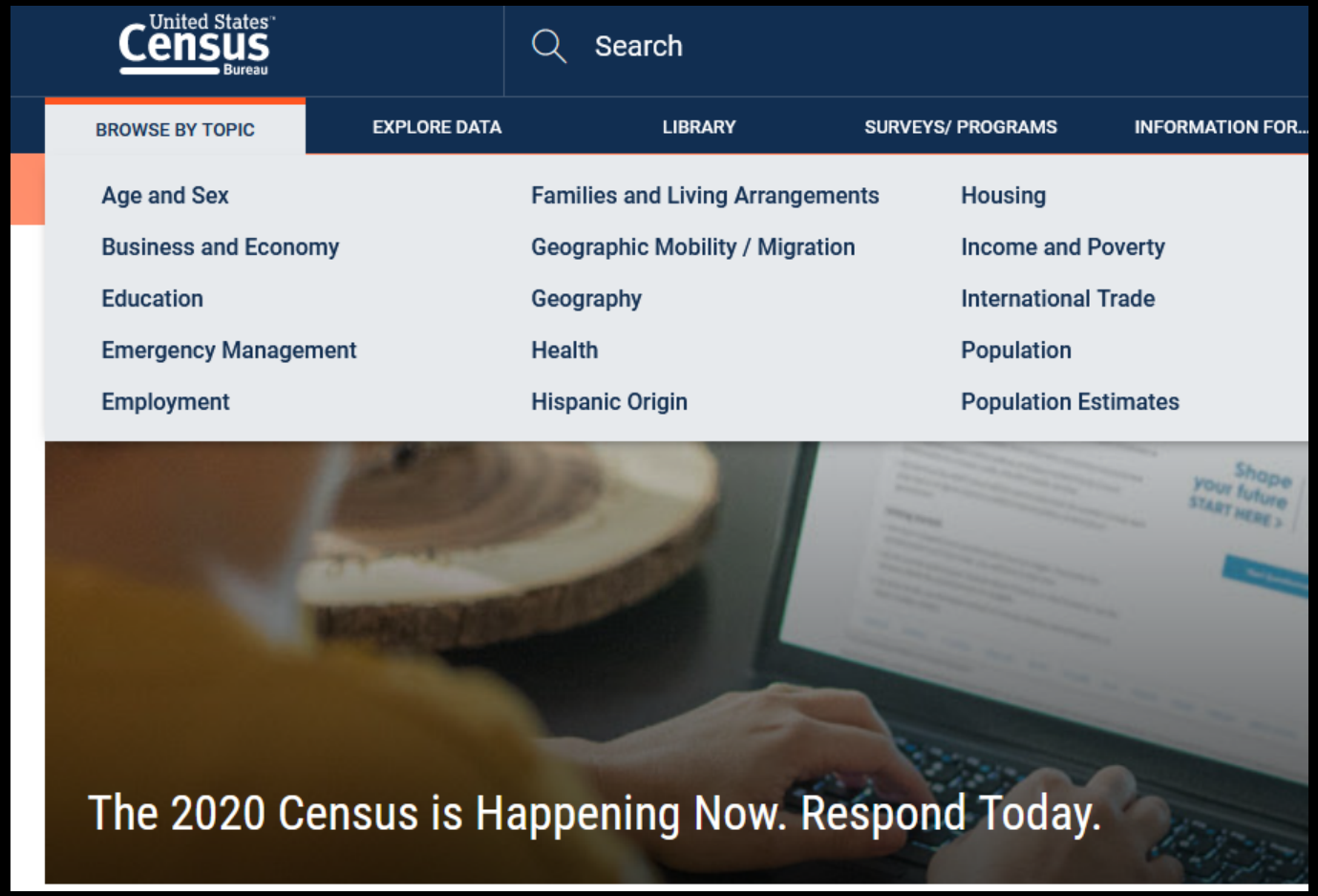
"*Between the dawn of civilization and 2003, we only created five exabytes of information; now we are creating that amount every two days*."
(Hal Varian, Chief Economist with Google, cited in Kitchen, 2014)

2. What kind of data out there today for urban planning?





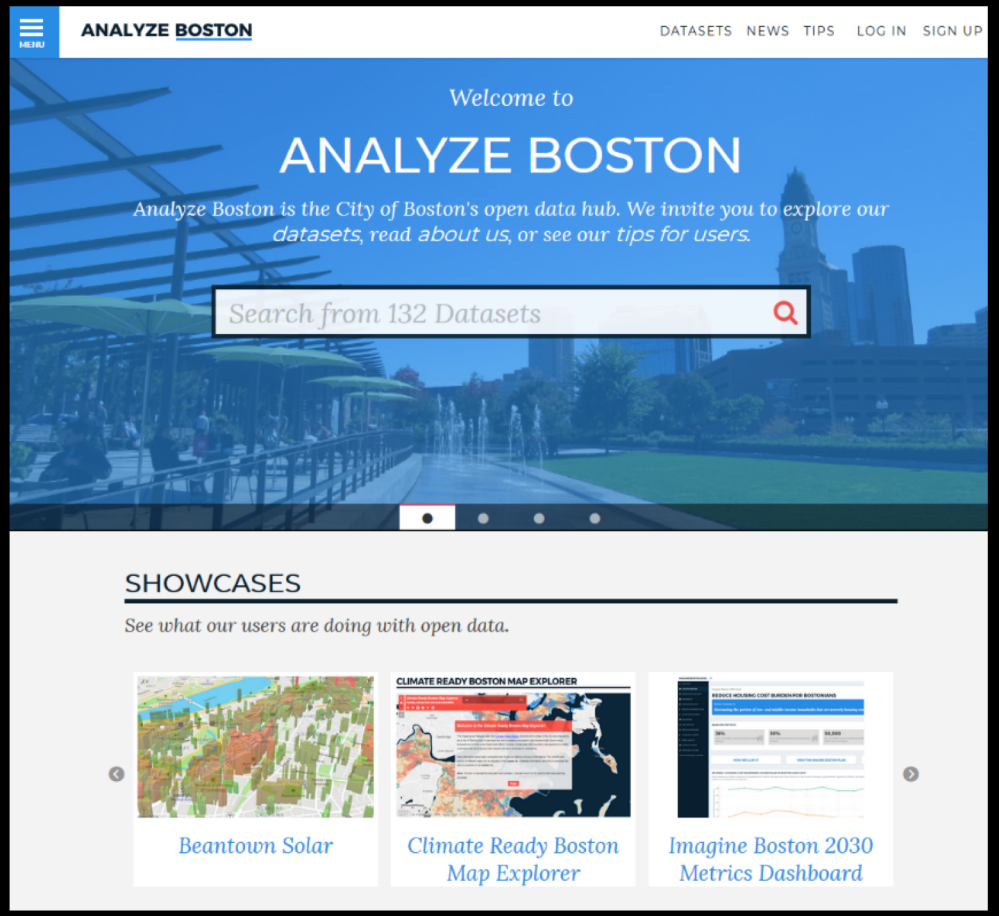
2. What kind of data out there today?



Captured directly through surveys, records

Derived through additional processing (e.g. poverty rate in a neighborhood)

2. What kind of data out there today?



Captured directly
through records, apps (311), etc..

Some are produced by residents.

2. What kind of data out there today?

Produced by the Information and communication technologies (ICTs) – digital and mobile devices and cloud computing, social media, the IoT (the internet of things)

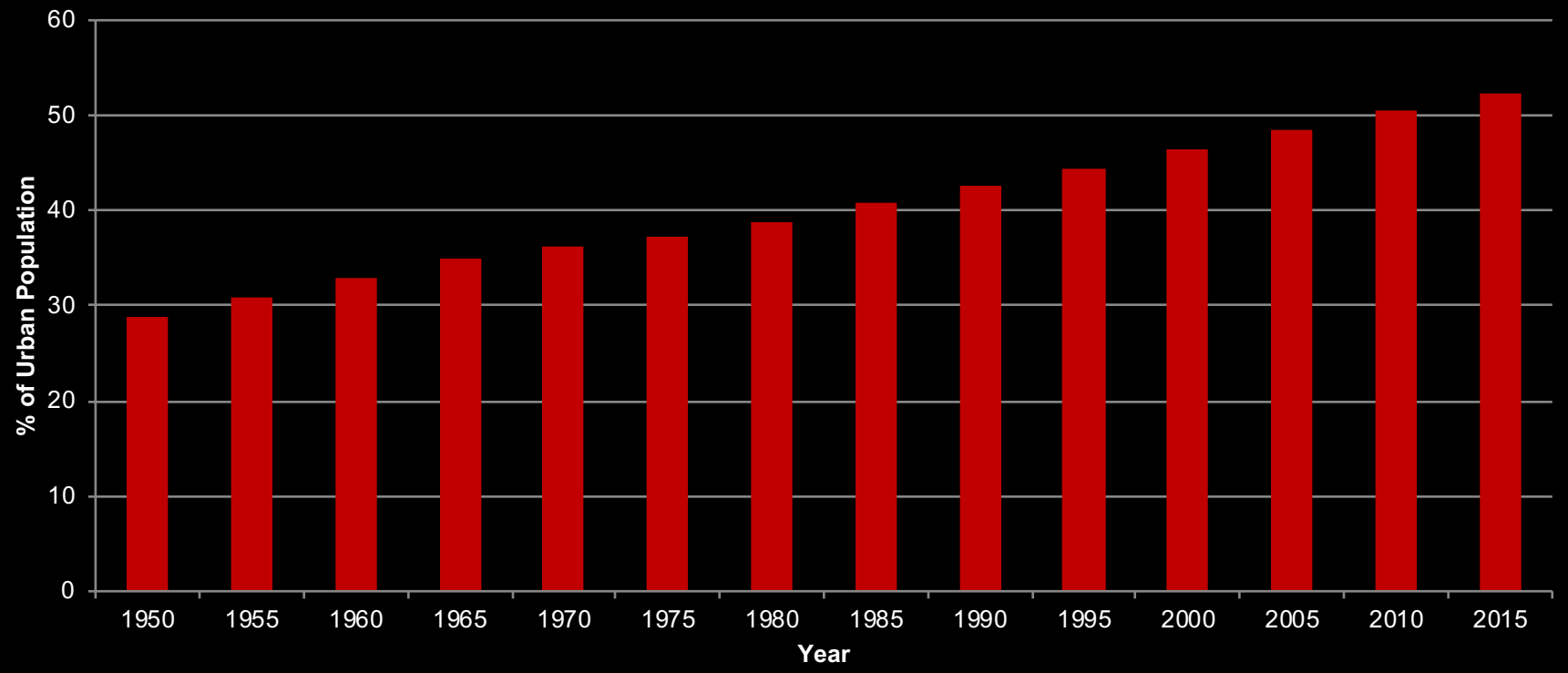
Data is now everyone's responsibility!

Image source: The Economist, March 23, 2016

Topic	Day	Week	Month	QTR
311 CALL CENTER PERFORMANCE	0.92	0.86	0.85	0.87
CODE ENFORCEMENT ON-TIME %	1.25	1.25	1.25	1.24
CODE ENFORCEMENT TRASH COLLECTION	1.25	1.22	1.18	1.19
GRAFFITI ON-TIME %		0.63	0.75	0.69
MISSED TRASH ON-TIME %	1.22	1.22	1.21	1.21
PARKS MAINTENANCE ON-TIME %	0.67	1.01	0.92	1.01
POTHOLE ON-TIME %	1.25	0.59	0.99	1.00
SIGN INSTALLATION ON-TIME %	1.00	0.90	0.87	0.90
SIGNAL REPAIR ON-TIME %	1.25	1.05	1.18	1.08
STREETLIGHT ON-TIME %	0.36	0.38	0.51	0.68
TREE MAINTENANCE ON-TIME %	1.25	1.25	1.20	1.21
ON-TIME PERMIT REVIEWS	0.97	0.95	0.90	0.90
LIBRARY USERS	1.25	1.15	1.15	1.15
BPS ATTENDANCE				1.04
BFD RESPONSE TIME	0.98	0.97	0.99	0.99

https://www.boston.gov/innovation-and-technology/cityscore

3. Why context is as important as the data itself?

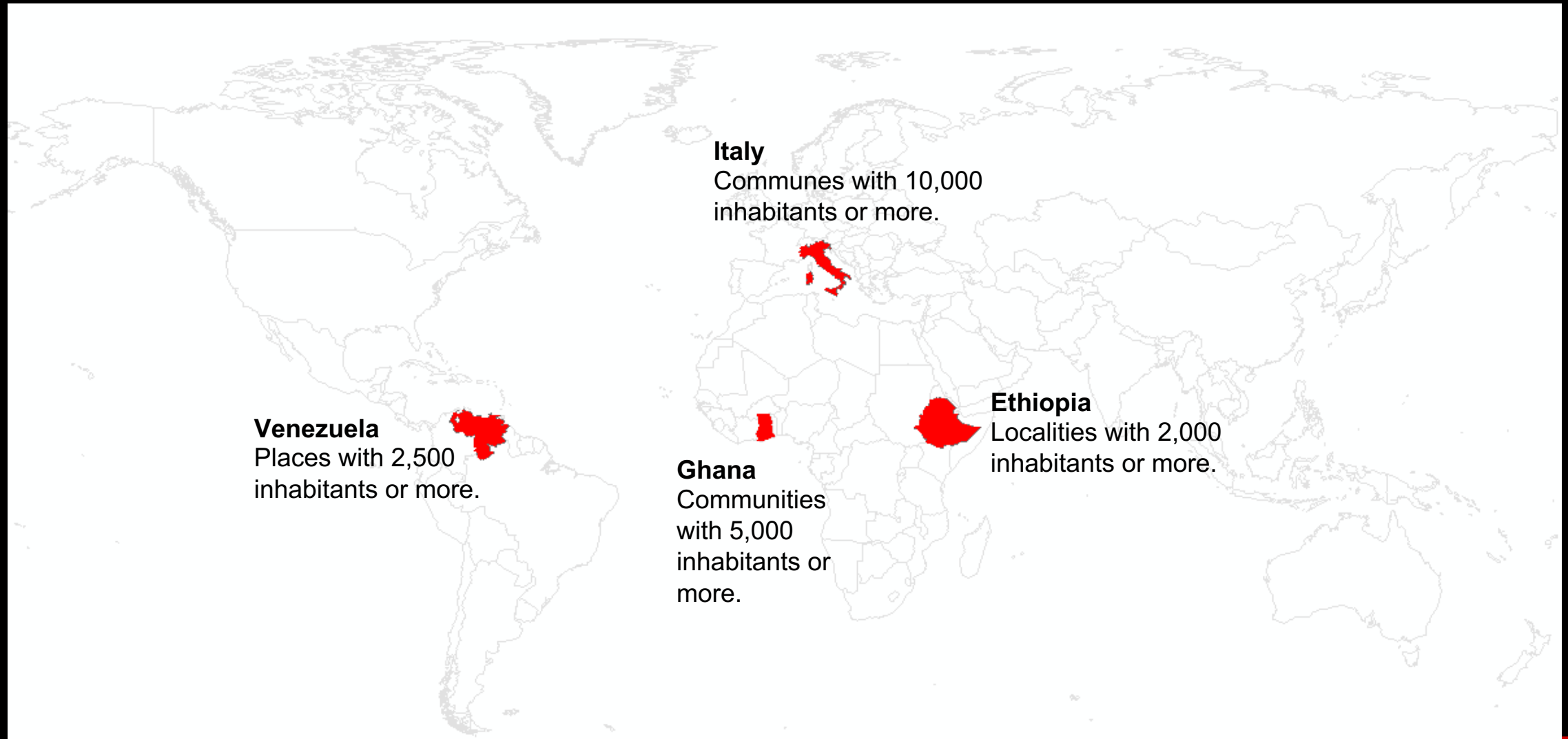


Source: The United Nations, 2017

True or False? There is a common global definition of **urban** so that we can study and measure it.

a. True
b. False

Data in Context



Data in Context

Canada
Places of 1000 or more, having a population density of 400 or more per sq. km

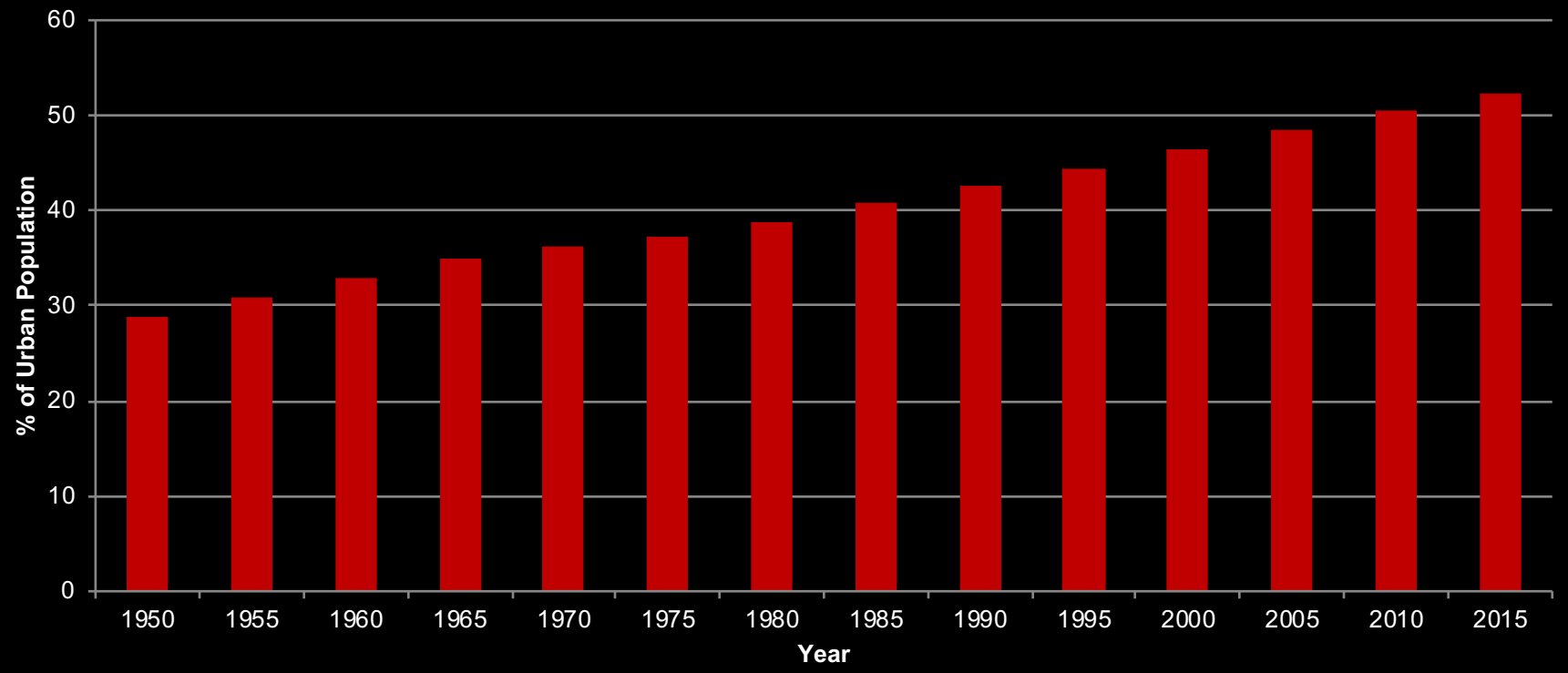
Iceland
Localities of 200 more and residences less than 200m apart

Japan
Areas 50,000 or more with 60% or more of the houses located in the main built-up areas and 60% or more of population engaged in manufacturing, trade, or other urban type of business.

Malaysia
Areas with a combined population of 10,000 or more. At least 60 per cent of their population (aged 10 years and over) engaged in non-agricultural activities. Areas with modern toilet facilities in their housing units.

Botswana
Agglomeration of 5000 or more inhabitants where 75% of the economic activity is non-agricultural.

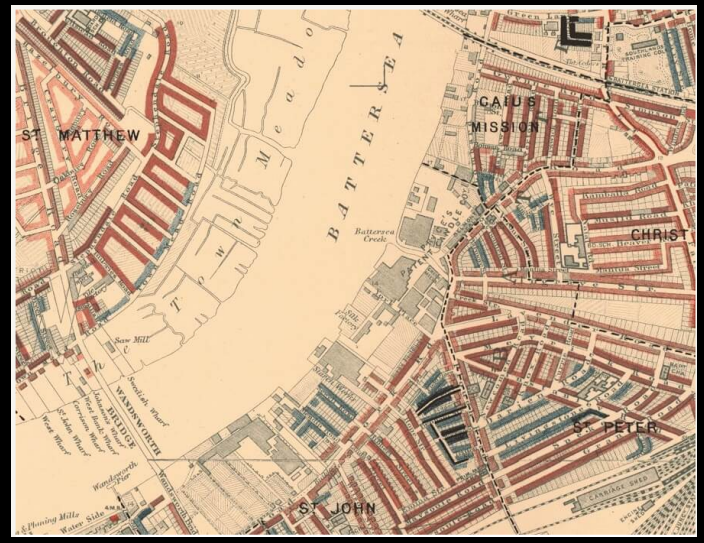
Why context is as important as the data itself?



Source: The United Nations, 2017

Data

- socially constructed.
- ideologically loaded.
- public good (Open data portals)
- commodity to be traded (some ICT data).
- Socio-technical production.



Classification	Colour	
Lowest class. Vicious, semi-criminal.	Black	
Very poor, casual. Chronic want.	Dark blue	
Poor. 18s. to 21s. a week for a moderate family.	Light blue	
Mixed. Some comfortable others poor.	Purple	
Fairly comfortable. Good ordinary earnings.	Pink	
Middle class. Well-to-do.	Red	
Upper-middle and upper classes. Wealthy.	Yellow	

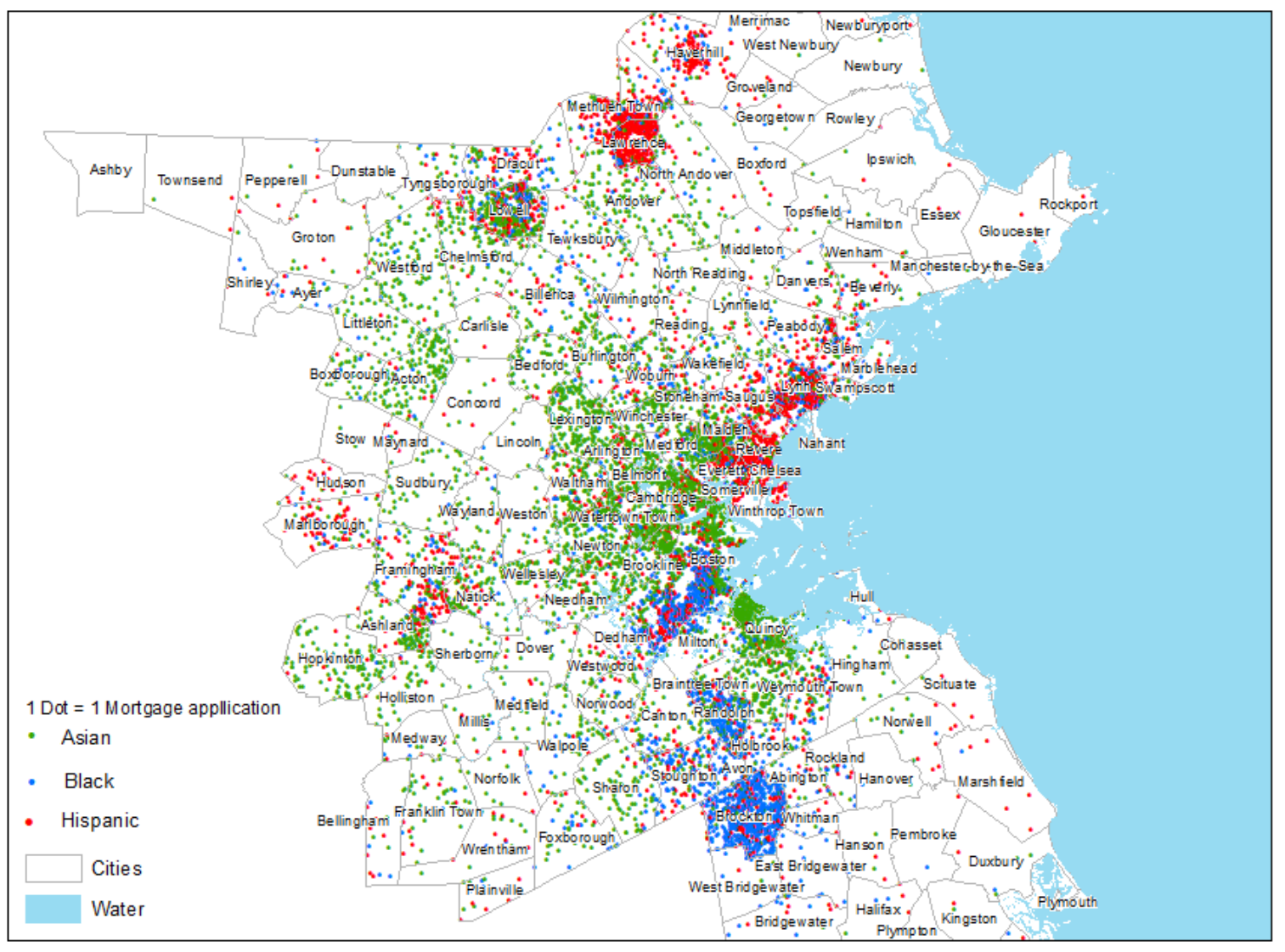
4. What do we do with data?

Wide spread ICT solutions are energy management systems, traffic control systems, smart grids, urban data platforms.

CityScore, Boston

Topic	Day	Week	Month	QTR
311 CALL CENTER PERFORMANCE	0.92	0.86	0.85	0.87
CODE ENFORCEMENT ON-TIME %	1.25	1.25	1.25	1.24
CODE ENFORCEMENT TRASH COLLECTION	1.25	1.22	1.18	1.19
GRAFFITI ON-TIME %		0.63	0.75	0.69
MISSED TRASH ON-TIME %	1.22	1.22	1.21	1.21
PARKS MAINTENANCE ON-TIME %	0.67	1.01	0.92	1.01
POTHOLE ON-TIME %	1.25	0.59	0.99	1.00
SIGN INSTALLATION ON-TIME %	1.00	0.90	0.87	0.90
SIGNAL REPAIR ON-TIME %	1.25	1.05	1.18	1.08
STREETLIGHT ON-TIME %	0.36	0.38	0.51	0.68
TREE MAINTENANCE ON-TIME %	1.25	1.25	1.20	1.21
ON-TIME PERMIT REVIEWS	0.97	0.95	0.90	0.90
LIBRARY USERS	1.25	1.15	1.15	1.15
BPS ATTENDANCE				1.04
BFD RESPONSE TIME	0.98	0.97	0.99	0.99

How is it put to work to improve people's lives?



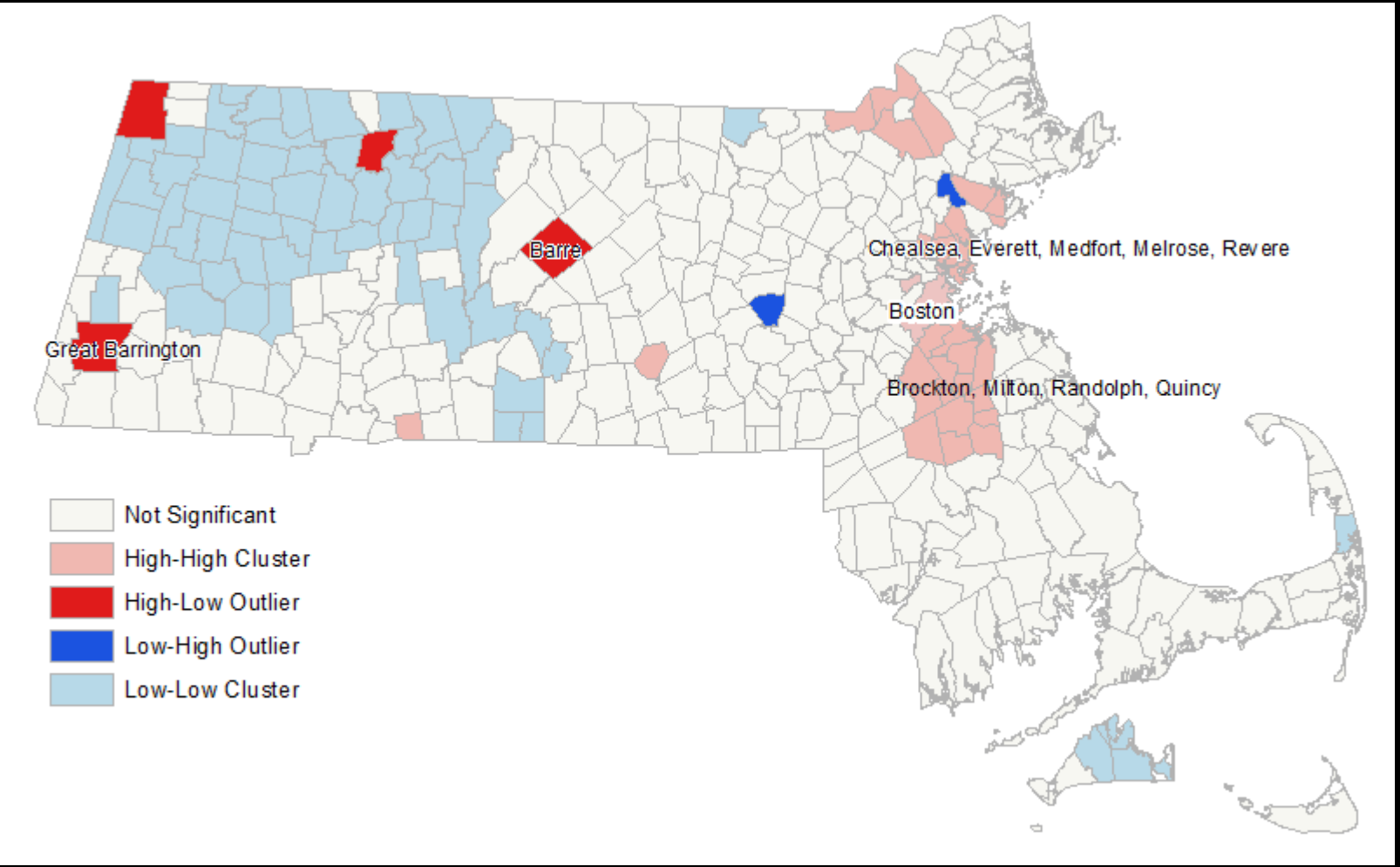
Boston is one of the limited number of cities in Greater Boston where Blacks or African American apply for home purchase loans.

$$I_i = \frac{x_i - \bar{X}}{S_i^2} \sum_{j=1, j \neq i}^{n} w_{i,j}(x_j - \bar{X})$$

$$S_i^2 = \frac{\sum_{j=1, j \neq i}^{n} (x_j - \bar{X})^2}{n - 1}$$

"Everything is related to everything else, but near things are more related than distant things."
(Waldo Tobler, 1970)

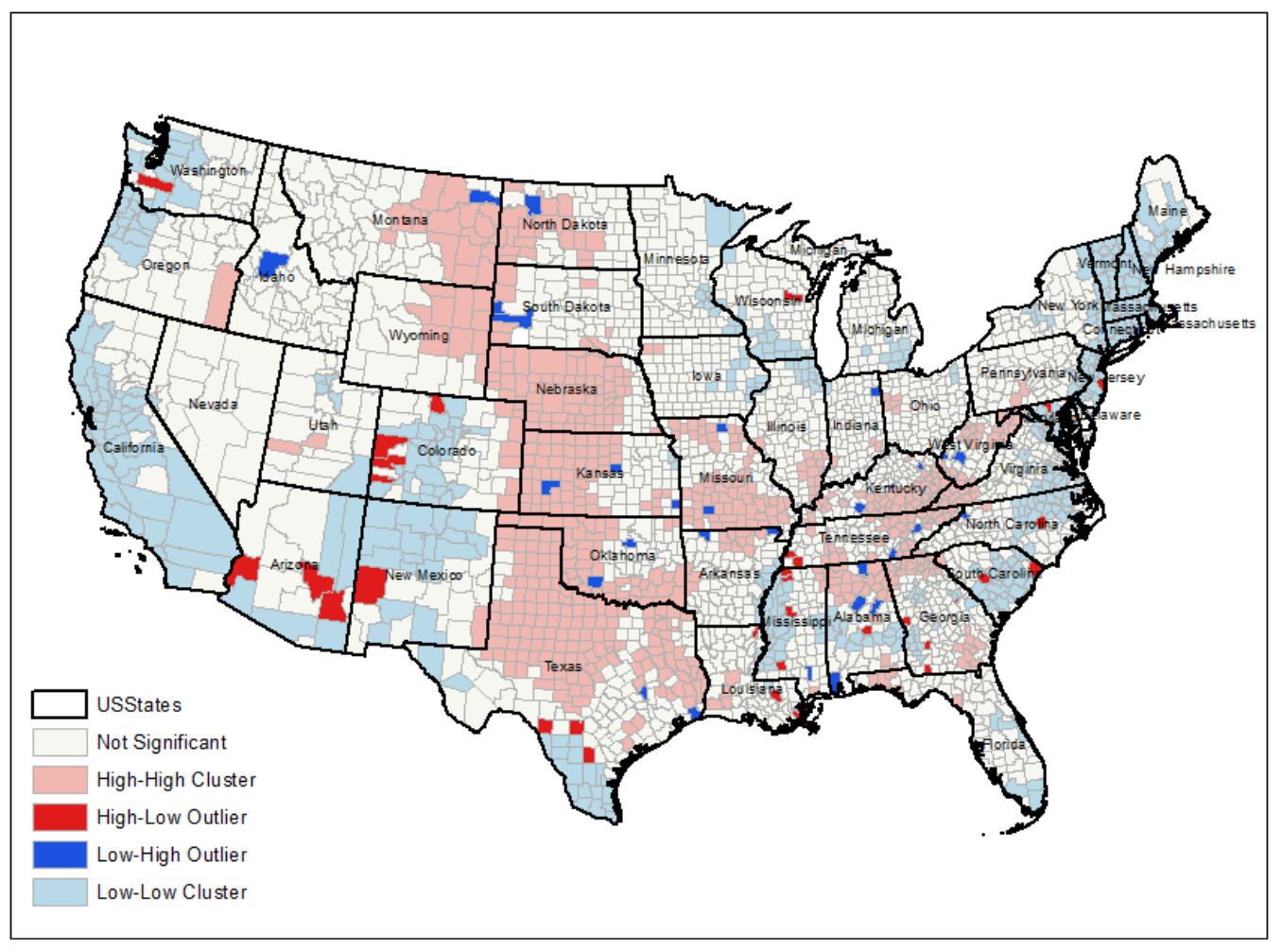
COVID-19 Hotspots and Coldspots in MA



Pink areas are 'hot' groups of towns/cities significantly affected by the virus. It is, collection of towns/cities, for instance, with a relatively high indicator which are also surrounded by areas of high indicators.

Outliers are areas that have a high number of cases within the town/city and fewer cases in neighboring towns/cities, highlighting an emerging risk or priority for containment.

Hot spots, Cold spots and Outliers in 2016 Presidential Election – Votes for the Republican Party



Pink areas are clusters that have a high percentage of votes for the Republican Party. It is, collection of counties with a relatively high indicator which are also surrounded by areas of high indicators.

CITIES & ANALYTICS: MAKING SENSE OF DATA

Monday 6 -- 8:45 pm

Yesim Sungu-Eryilmaz
yesims@bu.edu

What are the prerequisites of this course?

- Being passionate about urban planning and policy problems.
- Have an open and curious mind to learn about data and basic analytics.

#humansofcityplanning. #letsbereal #BUCPUA

THE *NEW* POSSIBLE

the post-pandemic BU CITY series

Short talks + Q&A w/ industry experts

Hosted by CITY PLANNING & URBAN AFFAIRS, MET COLLEGE
BOSTON UNIVERSITY

MEGA-URBANIZATION & SUSTAINABILITY
The case of Africa
WED, AUGUST 26, 6PM

THE CHANGING FACE OF TRANSPORTATION
WED, AUGUST 5, 6PM

GLOBALIZATION: PANDEMICS & PLANNING
FRI, AUGUST 14, 11:30AM

DATA & CITIES
WED, AUGUST 19, 6PM

PEOPLE PLACE & PRACTICE
Come join the movement!
FRI, July 31, 11:30AM

RSVP + More Info
All welcome. Free participation
http://www.bu.edu/cityplanning/thenewpossiblecity

BOSTON UNIVERSITY