SHARE IT! COMMUNICATING DATA INSIGHTS

James Madison, the fourth US president and renowned Father of the Bill of Rights, wrote: "A popular government, without popular information, or the means of acquiring it, is but a Prologue to a Farce or a Tragedy; or perhaps both. Knowledge will forever govern ignorance: And a people who mean to be their own Governors, must arm themselves with the power which knowledge gives."[1] Madison knew knowledge must be shared with the public in order for that public to make educated decisions in a democracy.

Sharing data does so much more than provide access to information. It creates trusting relationships, changes power dynamics, teaches us about policies, fosters debate, and helps to generate collaborative knowledge sharing, all of which are essential to building strong, deliberative communities. Yet interpreting data in its "raw" form can be difficult for the average person; most people do not have the skills to parse through data whether big or small. It follows that the insights found in data must be communicated so that anyone can understand them. How data is delivered affects how it performs in society.

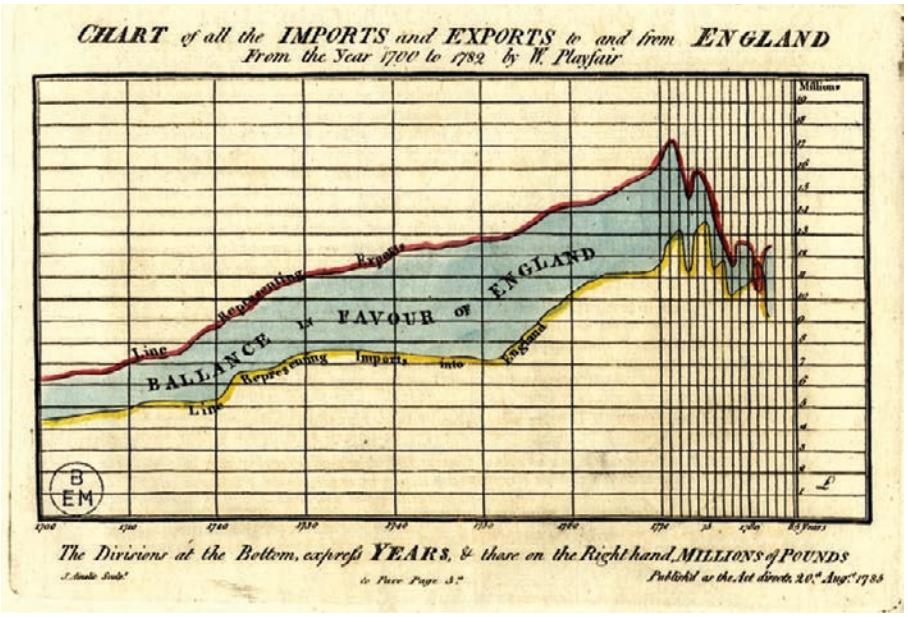
Sharing data is part of our everyday life. Whether that sharing is intentional or not, given or received, it transforms how we understand and *see* the world. We can all probably think of ways that data shared with us changed the way we looked at a topic or idea. Advertisers know this well; they often employ data to sell us everything from the latest fad diet to the most effective toothpaste. In fact, data is used as evidence to drive most anything, including policy. What's so powerful about using data to sell ideas is its easy

and seamless association with fact, which gives it legitimacy. That is why data's power needs to be wielded carefully: there is a fine line between using data to improve society and doing it harm.

In this chapter we look at the various ways that sharing data can provide insights to create policy action—focusing largely on sharing that happens through design communication such as data visualizations and mapping. The examples presented show that how we share data inherently constructs power relationships between those who collect, analyze, and present that data and those who consume the results. This chapter calls not only for governments to post their data online but also for data specialists everywhere to develop ways to share and communicate their results so that broad audiences can understand their work. Multimedia, visualizations, images, videos, art, performances, or interactive web pages help non-specialists to read the complexity of data without asking everyone to be a data scientist.

It is hard to believe that data visualizations such as pie charts, graphs, and histograms are a relatively new invention, when compared with maps. They appear on the scene in the eighteenth century and instantly become essential to communicating data. It is said that William Playfair's *Commercial and Political Atlas and Statistical Breviary* (1786) (figure 4.1) was the first of its kind, composed of graphs and pie charts, not maps that summarized the trade between England and other countries. Previously this sort of data had been printed as text. In a recent reprint of the Playfair *Atlas*, the editors discuss the brilliance of the development of Playfair's visualizations: "Graphs convey comparative information in ways that no tables of numbers or written accounts ever could. Trends, differences, and associations are seen at the blink of an eye. The eye perceives instantly what the brain would take second or minutes to infer from the table of numbers, and this is what makes graphs so attractive to scientists, business persons, and many others."[2] Playfair knew well that the tables he developed conferred a sense of legitimacy upon the message, and he used them to argue his position.[3]

Our media culture suits the consumption of data that can be seen in the "blink of an eye," as Playfair's early representations allowed. Data visualizations are excellent for telling stories and generating policy debates because they are bite-size, consumable thoughts.[4] The consumption of news and information has shifted toward using small, discrete visual packages posted on social media sites and blogs.[5] Data visualizations fit social media well because they create visual snapshots of complex ideas. Sharing images



4.1 Chart from William Playfair, *Commercial and Political Atlas and Statistical Breviary* (1786). This visualization is considered the first graph, and Playfair used it to show England's economic strength.

on sites like Twitter, Facebook, and Instagram allows ideas to spread widely while also allowing users of these sites to start to debate the ideas they present.

When data is shared for a specific purpose or intent, it is extremely powerful. That is why sharing data can be so controversial. Opening up raw data for anyone to use is important for transparency among groups and is a symbol of trust. At the same time it holds numerous risks such as privacy concerns, liable errors, and the exposure of underlying social patterns, which the person sharing the data might not have considered. Data shared in visual forms, such as maps, charts, and graphs, hold similar risks as we saw in the first chapter of this book. Highlighting populations on the margins of society can signal them for oppression, as evident in many of the maps developed by technocratic planners. Yet data visualizations, such as John Snow's now-famous map of cholera, can also expose poor living conditions and help to create policies that can improve the public's quality of life (see figure 1.6).

It is important to remember that data is a medium to construct and convey ideas, just as a collection of words makes a story, or an artist who uses

paint provides an image of the world. Like words on a page or brushstrokes on a canvas, the message that is shared through data visualization represents the thoughts and ideas of the person who shares it. Whether it's sharing "raw" data through an open data portal, or synthesized into visuals or an interactive website, data is shared with us through the lens of the person, group, or organization that provides it. Even in 1786, Playfair knew well that data is never unbiased and that the public rarely critiques its imperfections, making it a good tool for persuasion.

Constructing Change with Data

When I was eleven, I walked into my social studies class and on the wall saw the Brookes Slave Ship Map (figure 4.2). The week's lesson was on the abolition of slavery. The map illustrated how many enslaved people could fit on a typical slave ship carrying them from Africa to the West Indies. In an initial second, the graphic depiction conveyed the horrors of the slave trade—and that was the point: Where did they go to the bathroom? Did they see the light of day? Did they lie in their excrement? What if one of them died? Who would the corpse be next to? For how many months? How were they given food, lying so close together? Developed by British abolitionists in 1788 as a tool to argue against the slave trade, the map depicted the new standards for slave passage set in England's Regulated Slave Trade Act of 1788. The new regulations were an attempt to control the numbers of slaves allowed on ships, by specifying a space of 6′ × 1′4″ for each man; 5′10″ × 1′4″ for each woman; and 5′ × 1′2″ for each child. These maps showed that such an upgrade in space was still quite devastating.[6]

The Brookes Slave Ship Map of 1788, as the image came to be known, was quite effective when used to advocate for the abolition of slavery, and was widely reprinted. There is no denying the atrocities once you've seen this data visualization. It creates an emotional response. Its impact comes from its artifice and the narrative that the British abolitionists created around it. The map was developed by the Plymouth Chapter of the Society for Effecting the Abolition of the Slave Trade, and when their London counterpart saw the map they took it upon themselves to distribute eight thousand copies to hang on the walls of pubs throughout the country. The map went viral and was printed in pamphlets, newspapers, media, books, magazines, and so on.[7] Spreading the message of the maps was the abolitionists' intent. They used

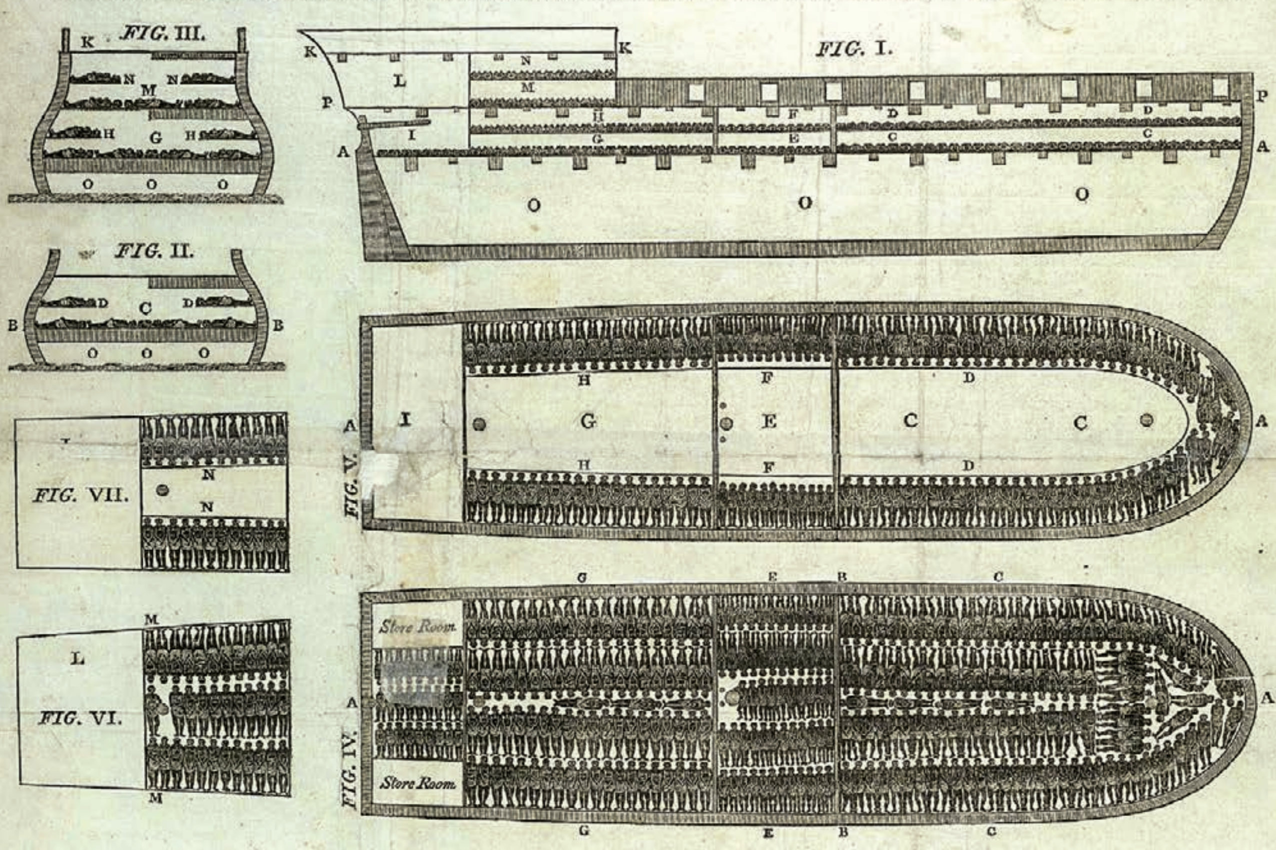
the maps as an awareness campaign about the horrors of slavery, and they were successful. The work was also included in Thomas Clarkson's *Abstract of Evidence*, a report delivered to the British House of Commons in 1790 and 1791, which helped to create the case for abolishing slavery.[8]

Although the image is said to be the most frequently used representation of the horrors of the slave trade, some groups find its use as an educational tool offensive. These critics see the map as minimizing the humanity of enslaved human beings, rendering them instead as inanimate objects lying and waiting to be saved by the inherently privileged white viewers, who were often complicit in the slavery system, and whom the maps were meant to persuade.[9] They argue that the design itself creates a kind of neutrality between the person viewing the map and what the map depicts, allowing viewers to remove themselves from the picture while allowing them to perceive themselves as saviors. I would argue, however, that this criticism is part of the map's artifice, and if it had been confrontational against the white majority, it might not have been as effective at communicating its message. Showing the horrors of the ship allowed those who viewed the map to interpret its message and create their own meaning from them.

I bring up the much-discussed Brookes Slave Ship map here because it demonstrates the intentionality of creating such images. It also illustrates how differently the image can be read, which helps explain why we must consider the potential ways that the data we share could be used to do harm (see chapter 3). There is no question the Brookes Slave Ship Map was an effective tool. It was made by relatively powerful people to persuade those in governmental power to abolish slavery, and it worked. The criticism of the map by those who see it as reinforcing dominant power narratives may be warranted. Are its critics arguing that the map should not be used? Absolutely not, but they are reminding us that we are not getting the whole story of how and why these maps were created. Their critiques remind us that even though the maps did do some good in the world, they also did so from a position of power.

How do we juxtapose the bias in maps and data visualizations with the good they were able to do? As the child seeing the map for the first time in my social studies class, it taught me a lot in one instant, but it was the narrative my teacher provided around the abolitionist movement that helped me to understand its power. Data visualizations help create a narrative around an idea, and it's the narrative that ultimately has the ability to change people's hearts and minds. When using data for action, we must focus on the story we

DESCRIPTION OF A SLAVE SHIP.



The PLAN and SECTIONS annexed exhibit a slave ship with the slaves stowed.‡ In order to give a representation of the trade against which no complaint of exaggeration could be brought by those concerned in it, the *Brooks* is here described, a ship well known in the trade, and the first mentioned in the report delivered to the House of Commons last year by Captain Parrey, who was sent to Liverpool by Government to take the dimensions of the ships employed in the African slave trade from that port. These plans and sections are on a scale of the 8th of an inch to a foot.

DIMENSIONS OF THE SHIP.

	Feet	Inches
Length of the *Lower Deck*, gratings, bulk-heads, included at AA	100	0
Breadth of *Beam* on the *Lower Deck* inside, BB	25	4
Depth of *Hold*, OOO from cieling to cieling	10	0
Height between decks from deck to deck	5	8
Length of the *Mens Room*, CC on the lower deck	46	0
Breadth of the *Mens Room*, CC on the lower deck	25	4
Length of the *Platform*, DD in the mens room	46	0
Breadth of the *Platforms* in mens rooms on each side	6	0
Length of the *Boys Room*, EE	13	9
Breadth of the *Boys Room*	25	0
Breadth of *Platforms*, FF in boys room	6	0
Length of *Womens Room*, GG	28	6
Breadth of *Womens Room*	23	6
Length of *Platforms*, HH in womens room	28	6
Breadth of *Platforms* in womens room	6	0
Length of the *Gun Room*, II on the lower deck	10	6
Breadth of the *Gun Room* on the lower deck	12	0
Length of the *Quarter Deck*, KK	33	6
Breadth of *Quarter Deck*	19	6
Length of the *Cabin*, LL	14	0
Height of the *Cabin*	6	2
Length of the *Half Deck*, MM	16	6
Height of the *Half Deck*	6	2
Length of the *Platforms*, NN on the half deck	16	6
Breadth of the *Platforms* on the half deck	6	0
Upper deck, PP		

Nominal tonnage	297
Supposed tonnage by measurement	320
Number of seamen	45

The number of slaves which this vessel actually carried appears from the accounts given to Capt. Parrey by the slave-merchants themselves as follows:

Men	351	Total 609
Women	127	
Boys	90	
Girls	41	

The room allowed to each description of slaves in this plan is:

To the Men 6 feet by 1 foot 4 inches.
Women 5 feet 10 in. by 1 foot 4 in.
Boys 5 feet by 1 foot 2 in.
Girls 4 feet 6 in. by 1 foot.

‡ This is the usual manner of stowing the slaves, but it varies according to the policies of the ships, and the practice of different commanders.

With this allowance of room the utmost number that can be stowed in a vessel of the dimension of the *Brooks*, is as follows, (being the number exhibited in the plan) and is less than 1½ to a ton, viz. †

		On the Plan.	Actually carried.
Men—on the lower deck, at CC	124	190	351
Ditto on the platform of ditto, CC DD	66		
Boys—lower deck EE	46	70	90
Ditto—platform FF	24		
Women—lower deck, GG	83	183	127
Ditto—platform, HH	40		
Women Half deck, MM	36		
Ditto Platform ditto, NN	24		
Girls Gun room, II		27	41
General total		470	609

The principal difference is in the men. It must be observed, that the men, from whom only insurrections are to be feared, are kept continually in irons, and must be stowed in the room allotted for them, which is of a more secure construction than the rest.

In this ship the number of men actually carried was	351
The number of men stated in the plan at 1 foot 4 inches each	190
Difference	161

As the ship on this plan would stow 42 women boys and girls in the places here allotted them more than she did carry, supposing that number of men taken from the mens room, and placed in their stead, this will reduce the number of men to 309 in the mens room; of course the room allowed them, instead of being 16 inches as in the plan, was in reality only 10 inches each; but if the whole number 351 were stowed in the mens room, they had only 9 inches each to lay in.

The men therefore, instead of lying on their backs, were placed, as is usual, in full ships, on their sides, or on each other. In which last situation they are not unfrequently found dead in the morning.

The longitudinal section, fig. I. shews the manner in which the slaves were placed on all the decks and platforms, which is also further illustrated by the transverse sections, fig. II. & III. By which it appears, that the height between the decks is 5 feet 8 inches, which, allowing 9 inches for the platform and its bearers, makes the height between the decks and the platform 2 feet 9 inches; but the beams and their knees, with the carlings, taking 4 inches on an average, this space is unequally divided, and above or under the platforms cannot be estimated at more than 2 feet 7 inches; so that the slaves cannot, when placed either on or under the platform, relieve themselves by sitting up; the very short ones excepted, nor can they, except on board the larger vessels. The average of 9 vessels measured by Captain Parrey, being mostly large ships, was only 5 feet 2 inches. The height of the Venus between decks was 4 feet 2 inches; of the Kitty, 4 feet 4 inches, both of which had platforms. In these smaller vessels therefore they they have not 2 feet under or upon the platforms.

In fig. 1. under the upper deck PP, and the lower deck AA, the beams and the intervening carlings are represented by shaded squares. The beams are also introduced on one side of the transverse sections II and III, in order to shew the space which a slave placed under a beam has to lie and breathe in.

† It must be noted, that every possible advantage of stowing is allowed in the plan. There are or ought to be in each apartment two or more proper tubs; there are also stanchions to support the platforms and decks; for which no deduction is made; but the deck is supposed clear of every incumbrance whatever.

It may be expected, from this mode of packing a number of our fellow-creatures, used in their own country to a life of ease, and from the anguish of mind their situation must necessarily create, that many of them fall sick and die. Instances sometimes occur of horrible mortality. The average is not less than 1-5th, or 20 per cent. The half deck is sometimes appropriated for a sick birth; but the *men slaves* are seldom indulged the privilege of being placed there, till there is little hope of recovery. The slaves are never allowed the least bedding, *either* sick or well; but are stowed on the bare boards, from the friction of which, occasioned by the motion of the ship, and their chains, they are frequently much bruised; and in some cases the flesh is rubbed *off* their shoulders, elbows, and hips.

It may not be improper to add a short account of the mode of securing, airing, and exercising the slaves.

The women and children are not chained, but the men are constantly chained two and two; the right leg of one to the left leg of the other, and their hands are secured in the same manner.

They are brought up on the main deck every day, about eight o'clock, and as each pair ascend, a strong chain, fastened by ring-bolts to the deck, is passed through their shackles; a precaution absolutely necessary to prevent insurrections.—In this state, if the weather is favourable, they are permitted to remain about one-third part of the twenty-four hours, and during this interval they are fed, and their apartment below is cleaned; but when the weather is bad, even these indulgencies cannot be granted them, and they are only permitted to come up in small companies, of about ten at a time, to be fed, where after remaining a quarter of an hour, each mess is obliged to give place to the next in rotation.

In very bad weather, some are unavoidably brought on deck: there being no other method of getting water, provisions, &c. out of the hold, but by removing those slaves who lie on the hatch-ways. The consequence of this violent change from their rooms, which are inconceivably hot, to the wind and rain, is their being attacked with coughs, swellings of the glands of the neck, fevers, and dysenteries; which are communicated by infection to the other slaves, and also to the sailors.

The only exercise of the men slaves is their being made to jump in their chains; and this, by the friends of the trade, is called *dancing*.

To persons unacquainted with the mode of carrying on this system of trading in human flesh, these Plans and Sections will appear rather a fiction, than a real representation of a slave-ship. They will probably object, that there is no room for stowing cables, and such other utensils and stores as are usually placed between decks. In a slave ship (i. e. a full one) these articles are either deposited in the hold, or piled upon the upper deck; and from thence, in case of bad weather, or accidents, no small confusion is occasioned.—It may be also said, the slaves are placed so very close, that there is not room for the surgeon to visit and assist them: The fact is, that when the surgeon goes amongst them, he picks out his way as well as he can, by stepping between their legs. He frequently finds it to be impossible to afford them that relief which an humane man (and such there are even in this trade) would willingly give them. When attacked with fluxes, their situation is scarcely to be described. To give an instance, (as related by an eye-witness) as it serves to convey some idea, though a very faint one, of the sufferings of those unhappy beings whom we wantonly drag from their native country, and doom to perpetual labour and captivity: "Some wet and blowing weather having occasioned the port holes to "be shut, and the grating to be covered, fluxes and fevers among the "negroes ensued. While they were in this situation, my profession "requiring it, I frequently went down among them, till at length their "apartments became so extremely hot, as to be only sufferable for a very "short time. But the excessive heat was not the only thing that rendered "their situation intolerable. The deck, that is, the floor of their rooms, "was so covered with the blood and mucus which had proceeded from "them in consequence of the flux, that it resembled a slaughter-house. "It is not in the power of the human imagination to picture to itself "a situation more dreadful or disgusting. Numbers of the slaves had "fainted, they were carried upon deck, where several of them died, and "the rest were, with difficulty, restored. It had nearly proved fatal to "me also."*

Another objection which may be stated, is, that here no room is allowed for the sailors hammocks. In slave ships, while the slaves are on board, the sailors have no other lodging than the bare decks, or (in large ships) the tops. From this exposure, they often are wet for a long time together, the rains in those climates being frequent and extremely heavy. There is in wet weather a tarpawling placed over the gratings; if the sailors to shelter themselves creep under this, they are exposed to the nuisance and infectious effluvia which continually exhale from the slaves below.

It appeared from the evidence given by the slave merchants last year before the House of Commons, that the employment of the seamen, viz. boating up the rivers after the negroes, guarding them on board, cleaning the vessel, &c. is of a nature offensive and dangerous beyond that of seamen in other services, and that the small-pox, measles, flux, and other contagious disorders, are frequent on board these ships.

It is therefore falsely said by the well-wishers to this trade, that the suppression of it will destroy a great nursery for seamen, and annihilate a very considerable source of commercial profit.—The Rev. Mr. Clarkson, in his admirable treatise on the Impolicy of the Trade, has proved from the most incontestable authority, that so far from being a nursery, it has been constantly and regularly a grave for our seamen; *for that in this traffick only, a greater proportion of men perish in* ONE *year, than in all the other trades of Great Britain in* TWO *years.*

Besides the time spent on the coast to complete their cargoes, which sometimes lasts several months, the slaves are from six to eight weeks on their passage from thence to the West-Indies.

Now let any person reflect on the situation of a number of these devoted people, thus managed and thus crammed together, and he must think it dreadful, even under every favourable circumstance of an humane captain, an able surgeon, fine weather, and a short passage. But when to a long passage are added, inhuman treatment, scanty and bad provisions, and rough weather, their condition is miserable beyond description. So destructive is this traffick in some circumstances, particularly in bad weather, when the slaves are kept below, and the gratings covered with tarpawlings, that a schooner, which carried only 140 slaves, meeting with a gale of wind which lasted eighteen hours, no less than 50 slaves perished in that small space of time.

As then the inhumanity of this trade must be universally admitted and lamented, people would do well to consider that it does not often fall to the lot of individuals, to have an opportunity of performing so important a moral and religious duty, as that of endeavouring to put an end to a practice, which may, without exaggeration, be stiled one of the *greatest evils at this day existing upon the earth.*

* Falconbridge's Account of the Slave Trade, page 31.

LONDON: PRINTED BY JAMES PHILLIPS, GEORGE-YARD, LOMBARD-STREET. M,DCC,LXXXIX.

want to tell with the data. This is important to remember when we use data visualizations for policy change—and this point is central to Data Action.

4.2 Brookes Slave Ship Map, 1788. Source: "Diagram of a Slave Ship," 1801, https://www.bl.uk/learning/timeline/item106661.html.

Sharing Data Creates Trust and a Commitment to Place

Sharing data can create civic action, and this chapter's case study shows how sharing data openly, through both "raw" data and visualizations, can change power dynamics and provide an essential resource for communities. Our case study takes us to Nairobi, Kenya, where I have worked since 2006 to improve the conditions for mobility in the city and have successfully used data as a tool in that work. Probably the most important benefit of sharing data with the people of Nairobi is that it created a sense of trust and commitment to *place*. Data standards are very important when sharing data, and by sharing our transport data in a standardized transport format we helped extend its use beyond the project we were working on; so that others could easily pick it up and use it. Visualizing their city's data transformed the way Nairobians saw their transportation system, which comprises an informal network of buses called matatus, and it created a platform for discussion. The visuals became an instant planning tool, helping matatu owners, the government, foreign aid agencies, and the public as well to make more informed decisions about the future of Nairobi. Perhaps most importantly our team recognized that the data we collected in Nairobi represented the people who live there, and by including input from matatu drivers and owners as we developed our visualizations, we hoped the map would accurately represent them. How the map was presented to the public was also important; we did not intend the visualizations to be used to marginalize or criminalize certain forms of mobility but rather to allow them to be more integrated into planning processes.

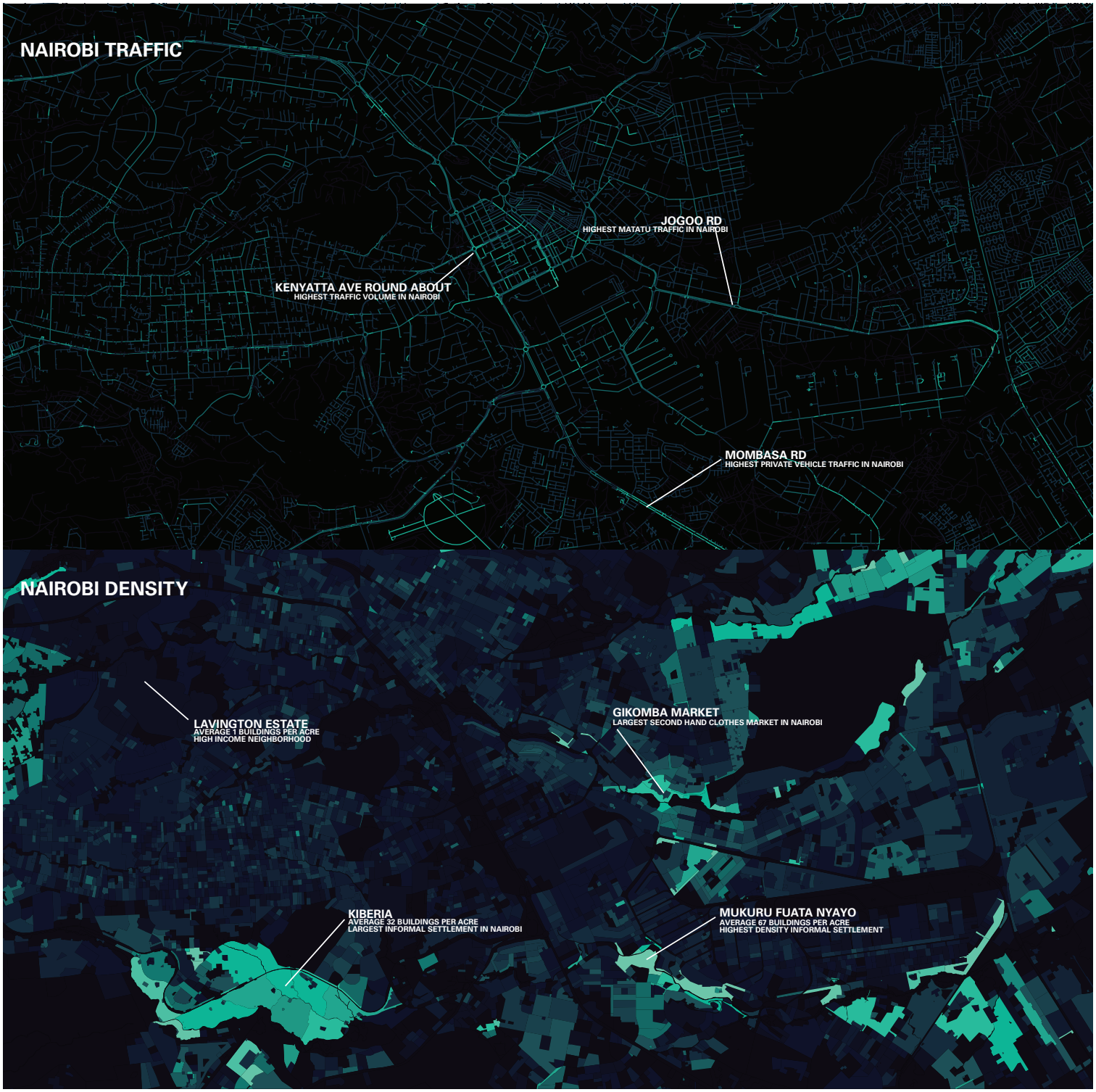
So, you might be asking, what is Nairobi's narrative around mobility? For starters, the city's biggest planning problem is congestion, which often brings the city to a standstill—it can take hours to travel a few miles. Understanding the root causes of congestion, as well as coming up with strategies to alleviate it, have been part of my work for some time. As early as 2006, I was involved in a project with Columbia University to develop a transportation model for Nairobi, the goal of which was to offer recommendations for ways to alleviate traffic congestion.

This is where our story begins. Obtaining data for our work in Nairobi was challenging, as is true for many rapidly developing cities. My team spent

months trying to acquire even the most basic data sets about roads, land use, and population. Even more disheartening was that this data about Nairobi existed thanks to earlier infrastructure projects, but my project team was unable to acquire it. This data, created by the Japan International Cooperation Agency (JICA) for Nairobi's National Spatial Data Infrastructure (NSDI) project, was meant to be openly available so it could be used to improve decision making in the city.[10] Our team struggled to obtain this open data, even though the funds that paid for its development specified that it should be made available via download on Kenya's Spatial Data Portal. A visit to the portal where the street files should have existed yielded a broken link. Eventually we determined the agency responsible for keeping the data; we were told that it would cost $10,000 for one license and that our multiple collaborators, including the Kenya Institute for Public Policy Research and Analysis (KIPPRA), as well as experts from our team at the University of California at Berkeley, would also have to pay. What we discovered was that data developed to be open to everyone was in fact a closed resource. Even though our project was an adequately funded study, we did not have the budget to acquire the information by paying for it.

Navigating the process to acquire data in Nairobi was hard. We seemed to encounter barriers at every turn. Even our local partner KIPPRA hesitated to share the basic survey data they obtained about transportation behaviors in Nairobi. There are many possible reasons for this: perhaps KIPPRA wanted to control the information so they could control the message; or maybe KIPPRA had become weary of the numerous outside actors who participated in the city's development, only to leave a few years later. They may have had concerns about our commitment to Nairobi. So many NGOs and multilaterals invest in projects in the city only to leave without a solid implementation plan. We had yet to create a trusting relationship to show our commitment to Nairobi.

While KIPPRA may not have trusted us, Nairobi's municipal government saw our interest in the data as an opportunity to extract payment. We had to come up with another way to acquire the data, and the best solution was to create it ourselves. Then we would share it for anyone to reuse, as we knew a lot of our colleagues in Kenya and beyond were experiencing the same problem. We enlisted the assistance of students from Columbia University and the University of Nairobi, and we created the first openly available digital data set of roads and land use by digitizing paper maps and satellite images (figure 4.3). This data set was essential for our transportation model; it helped



4.3 The first openly available data set for Nairobi, developed by Sarah Williams while Co-director of the Spatial Information Design Lab at Columbia University. Image created by Sarah Williams.

us make recommendations for removing roundabouts because the data models showed they were having severe, negative impacts on the organized flow of traffic.

For many years, the data we created for Nairobi was the only openly available road and land-use file. It was downloaded from our website hundreds of times and even integrated into Google Maps and Open Street Map—the roads seen in both mapping platforms were first derived from our original files. It's amazing that this data is still used today, even though it is quite dated, because it is the only openly available land-use data set for the city.[11] While the data itself was useful to analyze mobility, an even greater benefit was the trusting relationship it allowed us to develop with KIPPRA. Sharing data in this way manifested our commitment to transportation policy reform in Nairobi, and it was seen as a contribution beyond what was typical for outside actors. Sharing data helped create the trust needed to turn data into action.

Making and Sharing an Essential Resource to Create Action

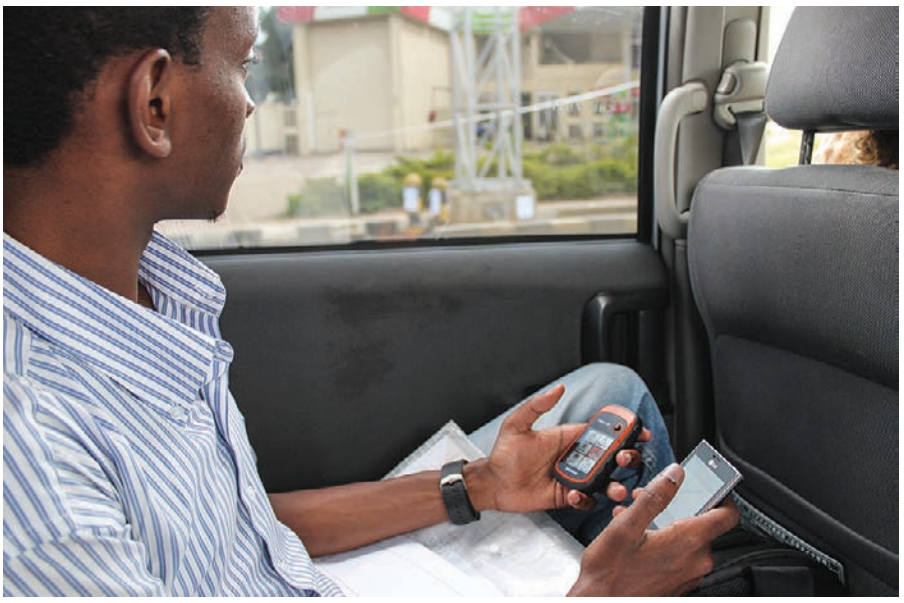
Even though our team created the first publicly available road data set for Nairobi, we were still missing essential data on Nairobi's matatu system—the main form of public transit in the city. An estimated 3.5 million Nairobi citizens depend upon this system every day, and yet at the time there was no publicly and widely available guide to navigating it. The matatu system is made up of privately owned, fourteen- to fifty-six-seat buses. These vehicles are owned by hundreds of different operators who self-regulate the system through the Matatu Owners Association, which helps settle disputes among owners and advocates. The system is loosely based on the former public bus system that fell to privatization in the 1980s. The matatu system sees little government involvement beyond providing licenses to drivers to operate—and from time to time the government creates enforceable mandates, such as banning matatus from the center of the city. Ultimately, however, the government does not want to be responsible for this system, nor to even appear to be in charge of it. Given these politics, the government has had little interest in collecting data about where the matatus go, yet having this information would allow for better transportation planning.

Therefore, our team decided to collect data on the matatu system—not only to provide an essential resource for our transport work but also

to create a resource that could be used beyond our project. The results had unprecedented impacts. We began in earnest with a visit to Nairobi's City Hall to identify the location of routes. We found a document that described in general terms some of the matatu routes: the #8 matatu, for example, travels from "Town to Kibera and Junction." The Microsoft Word document was often missing key information about the roads the matatus took along the routes, or where the matatus stopped to pick people up. It became clear that even the average Nairobi resident did not have information about where matatus traveled because there was no transport map of the city. As a result, my colleagues at the University of Nairobi's Center of Computing for Development Lab (C4D Lab), Columbia University's Center for Sustainable Development (CSUD), and the Civic Data Design Lab at MIT applied for a grant from the Rockefeller Foundation to collect geo-registered data on the location of matatu routes and stops that could be used for our model and eventually provide an essential map for both residents and visitors to Nairobi.

Armed with funding from the Rockefeller Foundation, this newly formed American–Kenyan team spent roughly the first eight months developing and testing cell phone software that would collect data while riding the matatus. During this development process, the research team in coordination with KIPPRA held a series of workshops for the local transit community. Workshop participants included members of the government, academia, the head of the Matatus Owners Association, matatu operators and drivers, nongovernmental organizations, and the local technology community. This collaboration succeeded in gaining support from these various actors (see figure 4.5).

Creating data in standardized formats is essential for extending its use beyond the research project for which it was intended. When we began collecting the data in early 2013, we used the General Transit Feed Specification (GTFS) open-data format to store the transit route data. GTFS is a data standard used for transit-routing applications, which allows software to find the best route when given a starting point and a destination. Most readers will be familiar with the Google Maps provision of driving, biking, and public transportation directions, all of which use the GTFS data standard to determine routes. There are numerous free applications that use GTFS data as their base, thereby extending the possibilities for what this particular collection of data can be used for upon release. There is, for example, an open-source application called Open Trip Planner, which is similar to Google

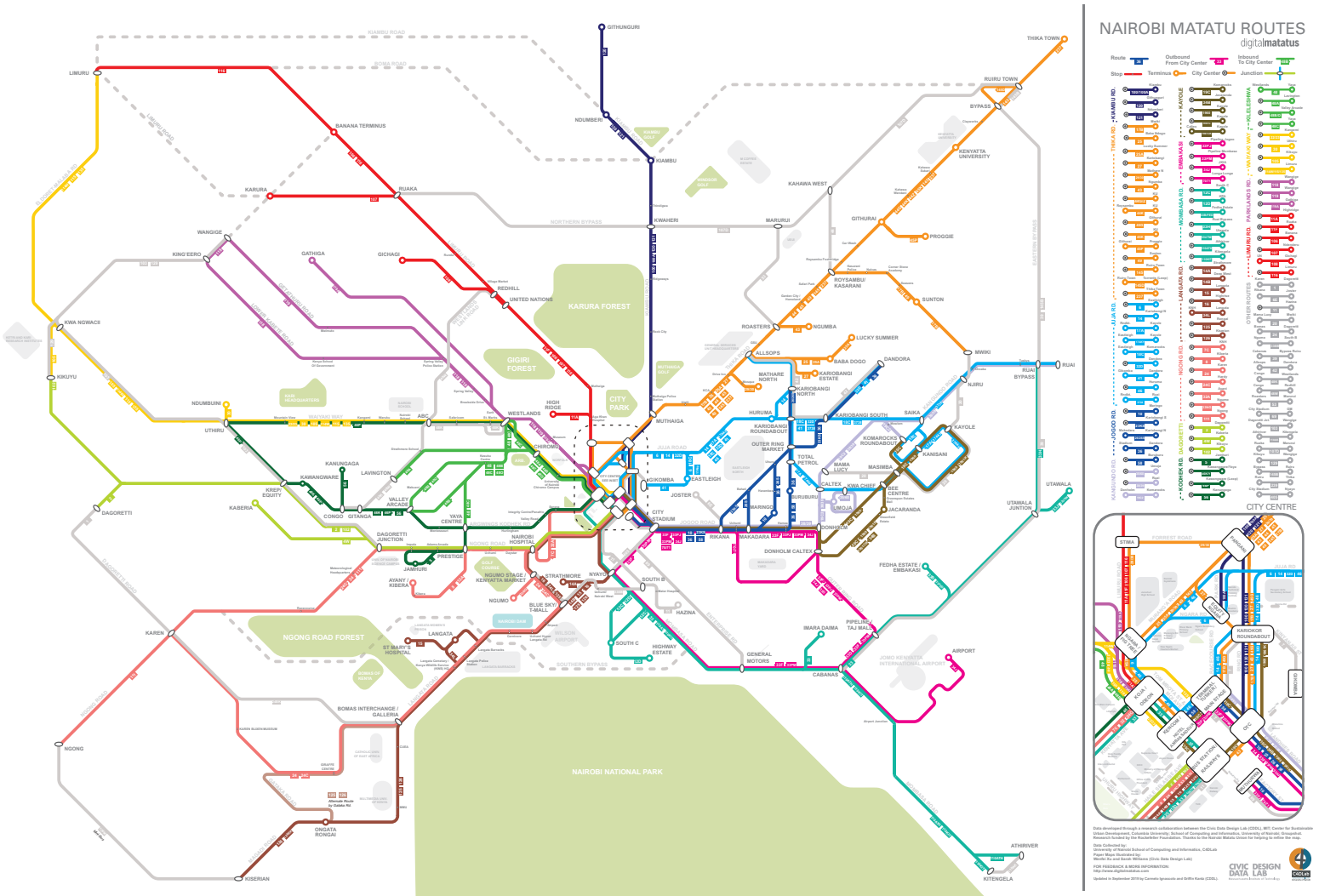


4.4a , 4.4b, and 4.4c University of Nairobi students collect data on the matatu system. Source: Elizabeth Resor and Adam White, courtesy of Digital Matatus

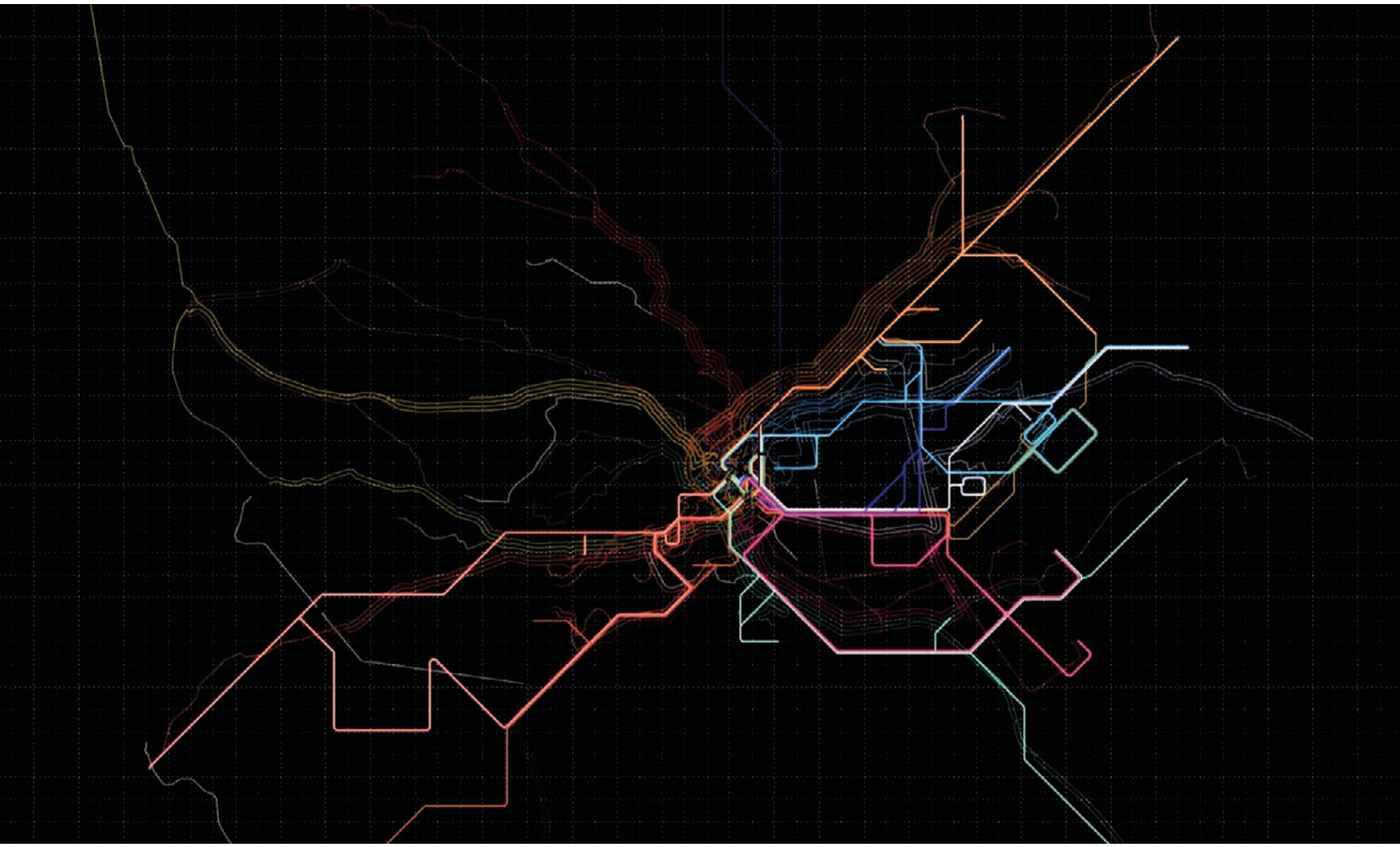
4.5 Workshop held to engage transportation stakeholders in Nairobi. Source: Sarah Williams

transit directions. But Open Trip Planner is a bit smarter because it has an analysis tool that allows users to determine how long it takes in both time and cost to get to different neighborhoods of the city, showing how accessible they are to transit.

Using GTFS, however, is no simple matter. It might be clear to a data scientist, but it is very abstract to the average person, including the government worker and the matatu driver. In order to make the data understandable and to communicate the matatu system to the broadest of publics, we decided to translate the data into a map that anyone could read and use.[12] Borrowing graphic language from London, Paris, and New York, we designed a stylized map of the Nairobi matatu system, grouping the matatus by the nearest major street and giving each corridor a color (figures 4.6 and 4.7). Major stops and points of interest including parks, airports, and landmarks were added to the map and allowed users to position themselves in Nairobi's geography.[13]



4.6 Final map of the Matatu System, 2014. Source: Digital Matatus



4.7 Process map for the development of the Digital Matatus System map. Source: Sarah Williams, Civic Data Design Lab

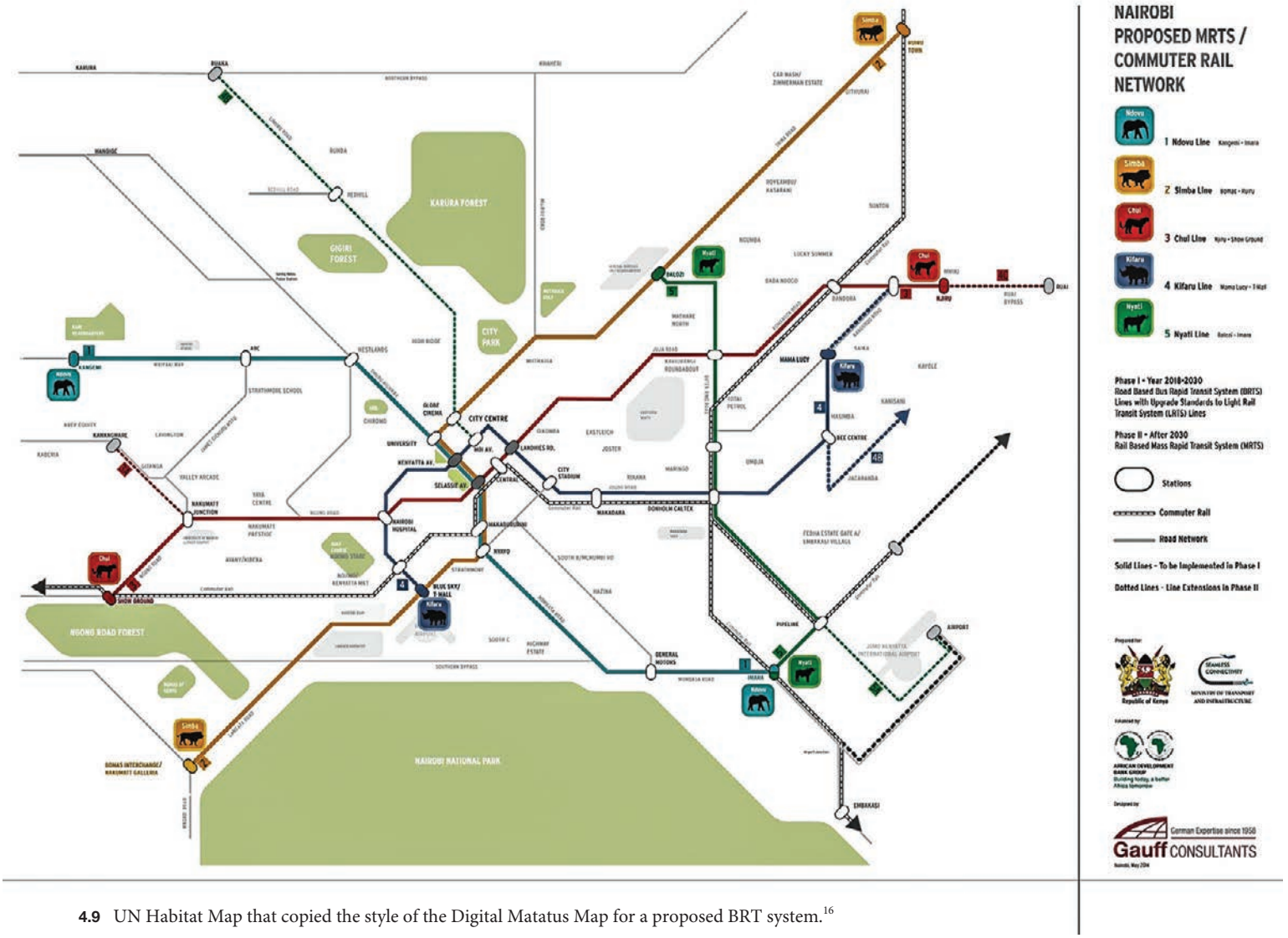
"For the First Time, We Can See Our Power"

One of the most powerful moments of the Digital Matatus project occurred when we were able to successfully share the data and maps we created with all of our partners (from government officials to matatu drivers) and the public of Nairobi. When the team presented the stylized maps to the matatu drivers and owners they were excited—for the first time they could see the comprehensive system they had created. The matatu owners are the de facto planners of the city's transportation system—and they instinctively began to use the map to plan new routes for Nairobi's future. The Digital Matatus map

also opened the eyes of government officials. Until they saw this map of matatu routes they were uninterested or perhaps didn't understand the need for our data collection project. Once visualized, the map became a powerful tool for their politics. The Ministry of Transport held a press conference releasing our visualization as the official matatu map of the city.[14] This happened because the government trusted the data we had created, and did so in large measure because we had kept officials informed about how the maps developed; all along the way, they trusted the data. The press conference helped create a discussion among the government, the transportation community, and the public about the future of the matatus system. Participants questioned the government about how it would respond to necessary service changes. After the press event, the downloadable maps went viral on social media. Large copies were printed in local newspapers allowing anyone to make use of the data we had collected and the map we had designed to visualize that data.

Our team also held a hackathon at the University of Nairobi to teach the local technology community about the GTFS data format and the open-source software that extends its use (figure 4.8). Our hackathon spawned two mobile applications using the collected data before the paper map was officially released. One of these programs, Ma3Route, became one of the most-used transit applications in Nairobi.[15] Users of Ma3Route share real-time data about the matatu system, noting changes to the routes, traffic

4.8 Hackathon at the University of Nairobi with the local technology community.



4.9 UN Habitat Map that copied the style of the Digital Matatus Map for a proposed BRT system.[16]

accidents, and traffic congestion. In due course, it won Kenya's Vision 2030 ICT Innovation Award in April 2014.

The Institute for Transportation and Development Policy (ITDP) and consultants at UN Habitat also used Digital Matatus data as the basis of the planning of Bus Rapid Transit (BRT) Service Plans for the city. Nairobi city planners used the data to develop a map of the BRT system that looked like a copy of our map. The city planners were smart to borrow the visual language of our map to earn support for the BRT project because the Digital Matatus map had become an icon in the city (figure 4.9).

4.10 The maps went viral on social media and the internet, and were printed in the local newspaper. This image shows the map centered on a page of Nairobi's newspaper *The Star*. Source: Photo by Sarah Williams.

Digital Matatus: Co-creation Is Powerful

The early work accomplished by KIPPRA paved the way for our now legendary Digital Matatus project because we created a trusting relationship between our research team and Nairobi's transport community. The work of digitizing the matatu routes encompassed many of the methods that I argue are important when employing data for action. Our team built a data set and developed open-source technology to create and edit the data. We shared the data openly by posting it online in a standardized format (to help extend its use beyond our Digital Matatus project) and by creating visualizations. In time, transportation actors in Nairobi used the data to help change some of their city's policies. Sharing data, both as visualizations and in a standardized data format, did indeed extend the life of our transport work in Nairobi and build trust. It allowed others to use the data for their own policy change, create a community around the data, and provide an essential resource to the public. The Digital Matatus project shows that data visualizations are

powerful vehicles for generating debate and presenting evidence for planning strategies. The stylized transit maps we developed allowed the government to engage in conversations with the public. Nairobi's matatu operators used the map to identify and develop new routes for the system. Perhaps more importantly, the citizens of Nairobi now have essential information for navigating their city.

Digital Matatus has inspired cities all over the world from Cairo to Bogotá—twenty-six cities in total at the time of this writing in 2019. This network led us to launch a Global Resource Center for the Development of Informal Transit Data with headquarters in Mexico City and Addis Ababa in 2018. The center provides open-source tools, trainings, links to other cities that have done this work, as well as assistance with policy impact and integration. Ultimately the Digital Matatus project has given life to a new form of collecting data on informal systems, one that is collaborative, open, and transparent.

Maps Are a Powerful Medium—They Persuade

Maps are associated with truths and can be powerfully persuasive; however, their presentation of the data can be intentionally or unintentionally misleading, and some people are therefore wary of using them. It can be easily argued that we rarely pay enough attention to the sources on the maps we read, let alone critique their accuracy or the data they are based on. Indeed, as we've seen, the very act of developing data visualizations involves a bias since designers must choose what to simplify and what to abstract to make their maps. Therefore, more critique is warranted.

Critical cartographers including the likes of Brian Harley, Denis Wood, John Pickles, Michael Curry, Jeremy W. Crampton, Sarah Elwood, Annette Kim, and Matthew Edney argue that maps are inherently political; that what is added to and left off a map illustrates or points to social constructs.[17] For example, cartographers often leave out poor areas or slums, alleyways, and much more to display a pristine representation of place that serves their purpose. Studying the symbols on a map can reveal systems of power and control. Critical cartographers ask us to interrogate the political meaning behind maps. Being political doesn't necessarily make the construct of a map harmful—quite the opposite. The politics of maps can be used for good, too.

David Harvey, a Marxist geographer, argues that maps are a tool of power and control because only those in power have the wherewithal to create them, and they often create a one-dimensional perspective on a topic or idea that helps them to retain that control.[18] Brian Harley, a well-known cartographic historian, argues that the practice of developing detailed property records such as cadastral maps is akin to the practices of old-world, agrarian landowners controlling their property. According to Harley, "Accurate, large-scale plans were a means by which land could be more efficiently exploited, by which rent rolls could be increased, and by which legal obligations could be enforced or tenures modified. Supplementing older, written surveys, the map served as a graphic inventory, a codification of information about ownership, tenancy, rentable values, cropping practice, and agricultural potential, enabling capitalist landowners to see their estates as a whole and better to control them."[19] Harley thought this description was fitting for city planners who, he believed, were using their data to divide the modern city into parcels that could be easily bought and controlled.[20] Rather than using the data to improve the quality of life for the citizens, he believed their data-collecting strategies used the cloak of scientific knowledge to advance predetermined objectives. Yet even with his cynicism toward the development of cadastral maps we read about in chapter 1, Harley is not arguing that we should not use maps, but rather that we ought to acknowledge that they are not "value-free."

Making a System Visible to Persuade

Some scholars might claim that by developing a map of Nairobi's matatu system we created a tool of oppression: we made visible an informal system, and the visibility could now be used as a tool to control the matatus with unnecessary regulation or perhaps even cause a complete dismantling of the system. Others might say that in developing a map in a style that mimics European and American representations of transportation, we were framing the data within a colonial construct. While these criticisms might be warranted, matatus have not been made illegal in Nairobi, and the government is still largely uninterested in regulating the vehicles and we have provided an essential resource that the public uses to navigate the city.

Let's take a moment to pick these ideas apart. Why weren't the maps in Nairobi used as a tool for control? I would argue this is because the maps and data were developed collaboratively, and through that effort the maps

were co-owned by multiple actors, with none having a singular power over its message. Nairobi's matatu drivers do equate the maps to those of European and American transport systems, and that gives them a sense of pride in their system. Rather than being viewed as part of an informal economy, equating the matatus with formal systems helps show that they contribute an important public infrastructure in the city.

The matatu maps "do no harm" because they tell the untold story of the matatu system. The narrative created by the matatu maps does not paint a picture of an unsafe, corrupt system that needs to be removed but rather one in which transportation actors need to work collaboratively to make the current system more usable. In other words the map using the European representation allowed the matatu driver to be heard. Just as the narrative around the Brookes Slave Ship Map was about abolishing slavery and was meant to speak to those that could change that narrative, the matatu map was to meant to speak to participants and stakeholders both local and foreign, showing that matatu drivers and owners currently run Nairobi's transport system and their voices should be heard. Previously, it was thought that each matatu owner was operating quite independently, and as such any planning decisions that addressed the whole system would be impossible. The Digital Matatus map was shown to those seeking to influence transportation policy in Nairobi, such as the World Bank, to demonstrate that working with the matatu association would be essential for any transportation plan for the city, including the integration of more formalized systems like the Bus Rapid Transit. It is important that we acknowledge maps as a tool for control, and the Digital Matatus project shows how we also used maps to control the narrative around matatus in Nairobi. Our team used the maps as an advocacy tool, representing our ideological positions toward helping those on the margins. There was certainly bias, but does that cause harm? This is the question all data visualizers must ask themselves.

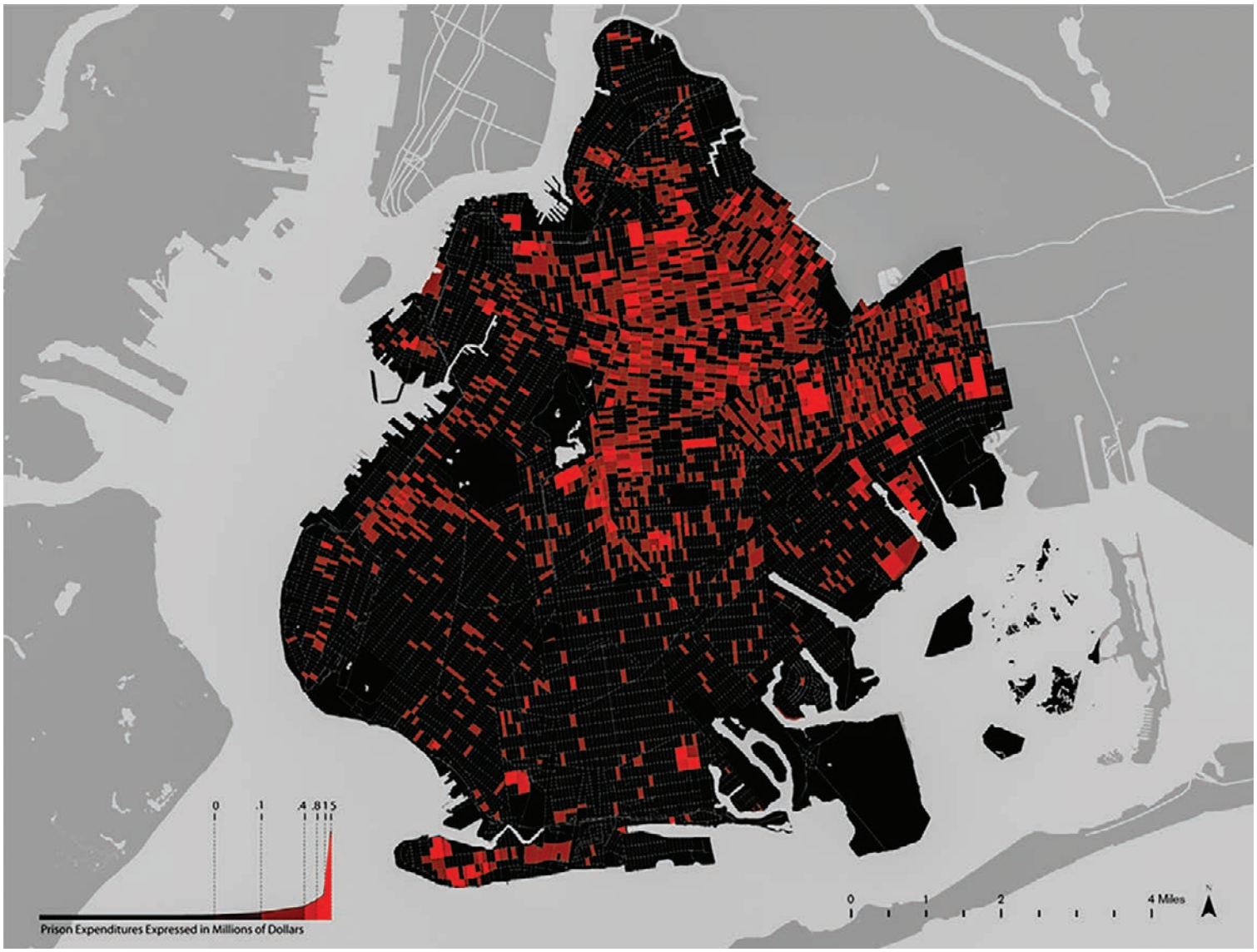
Transforming Map Power for Those on the Margins

The Million Dollar Blocks project transformed the idea of a map from one of power and control into a tool to advocate for those on the margins of society. Our team at the Justice Mapping Center and the Spatial Information Design Lab at Columbia University developed this project to create maps to expose the unconscionable cost of incarcerating people both in terms of actual dollars

spent and the cost to the communities the incarcerated come from. Our team transformed a typical crime map, which marks neighborhood hot spots in need of policing, to construct a map identifying neighborhoods in need of investment. Crime maps are often criticized because the way they are used does not address the systemic causes of crime, namely the social problems such as a structural racism, lack of proper access to education, healthcare, and jobs that stem from living in communities of poverty. So while crime might be reduced temporarily in one community because of a police crackdown, it often moves somewhere else. Million Dollar Blocks transforms the typical crime map from a tool of control into a tool to expose the massive amounts of money we spend to incarcerate people rather than address the systemic reasons for their incarceration.

The Million Dollar Blocks maps were developed using prisoner intake data, which includes information about where people lived before they went to prison. These addresses were plotted on a map and cross-referenced with the amount it cost to incarcerate these people, then data was summarized by totaling all the costs for each block. The result created a map showing that in some communities there are concentrations of blocks where more than a million dollars is spent to incarcerate residents (figure 4.11).[21] Often both physically and socially isolated, these neighborhoods typically lack the resources to alleviate the causes of incarceration, including access to education, job training programs, or prison reentry services.[22] The maps conveyed a simple idea: we spend millions of dollars to incarcerate people, so how might we spend that money better in these communities?

The images are at once alarming and captivating because they scale prison policy down to the size of a city block—something that everyone can understand (figure 4.13). Visualizing the data at the scale of the block was an important method for contextualizing the vast sums of money spent on incarceration. The maps show million-dollar blocks as bright red on a black background (a combination rarely seen on maps), which marks the issue as alarming at first glance. Maps showing race and poverty were presented alongside prison spending, allowing those who read the maps to connect incarceration to high levels of poverty and racial segregation. The causes underlying the million-dollar-block designation are multifaceted, of course, and there is no single answer. These maps were developed not to provide a solution, but rather to discuss the issue and present a message: we spend millions of dollars to incarcerate people in the United States, and the cycling of people in and out of prison has become a big business.[23]



4.11 This Million Dollar Blocks map shows prison expenditures in Brooklyn, NY in 2003. The bright red blocks show where over a million dollars was spent in just one year. *Source*: Maps created by Eric Cadora, Laura Kurgan, David Reinfurt, and Sarah Williams, courtesy of Spatial Information Design Lab.

4.12 This map, seen on the wall of the Design and the Elastic Mind Exhibition at MoMA in 2008, shows where each person incarcerated in 2006 lived and connects their home to the prison they were sent to in upstate New York. The blocks shown in bright red underneath the lines are ones where more than a million dollars was spent to incarcerate residents of that block. The maps on the top compare percentage of people of color, percentage of people in poverty, and prison admissions.
Source: Maps created by Eric Cadora, Laura Kurgan, David Reinfurt, and Sarah Williams, courtesy of Spatial Information Design Lab.

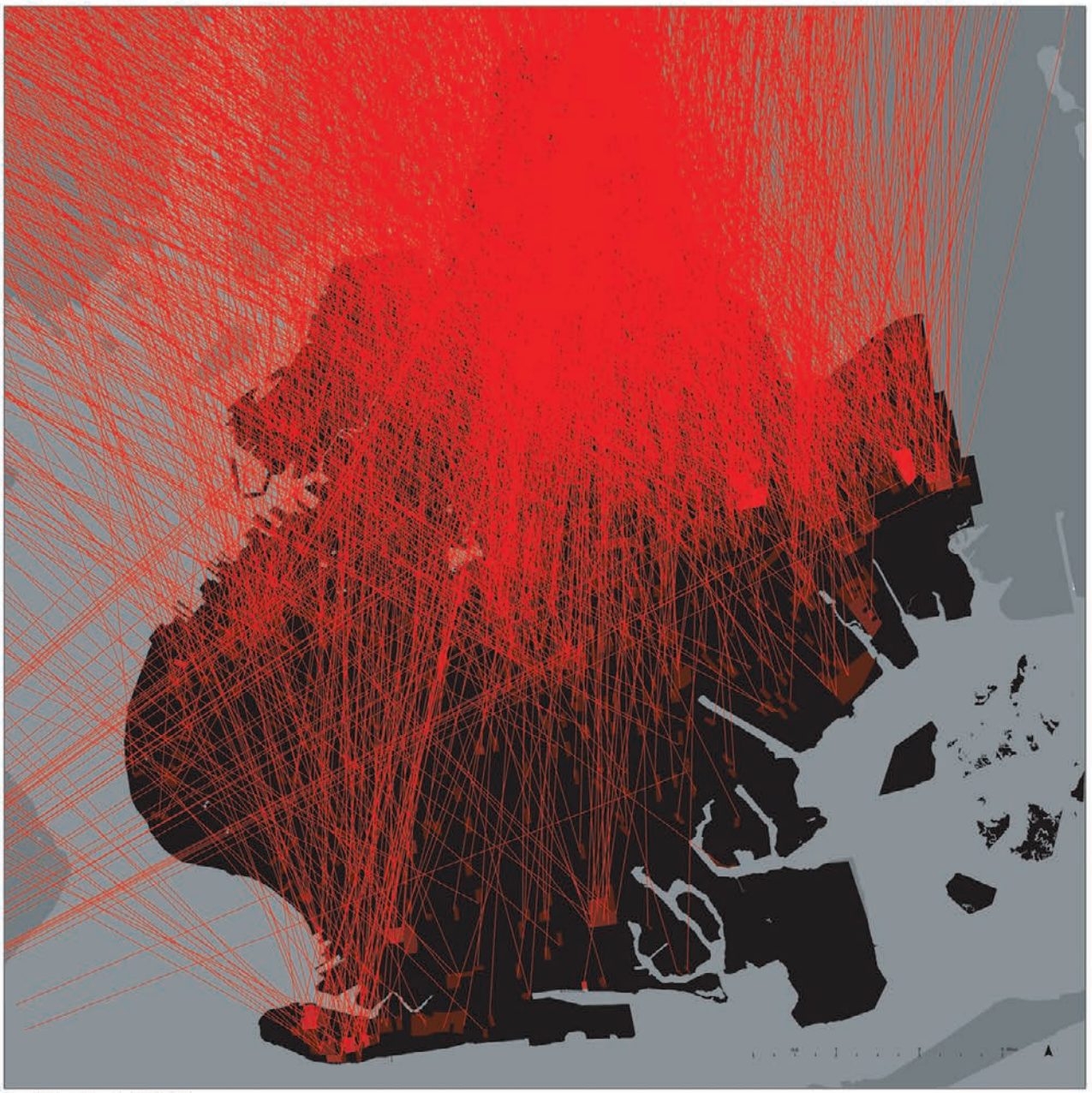
Percent Persons of Color, 2000.

Percent Persons Below Poverty Level, 2000.

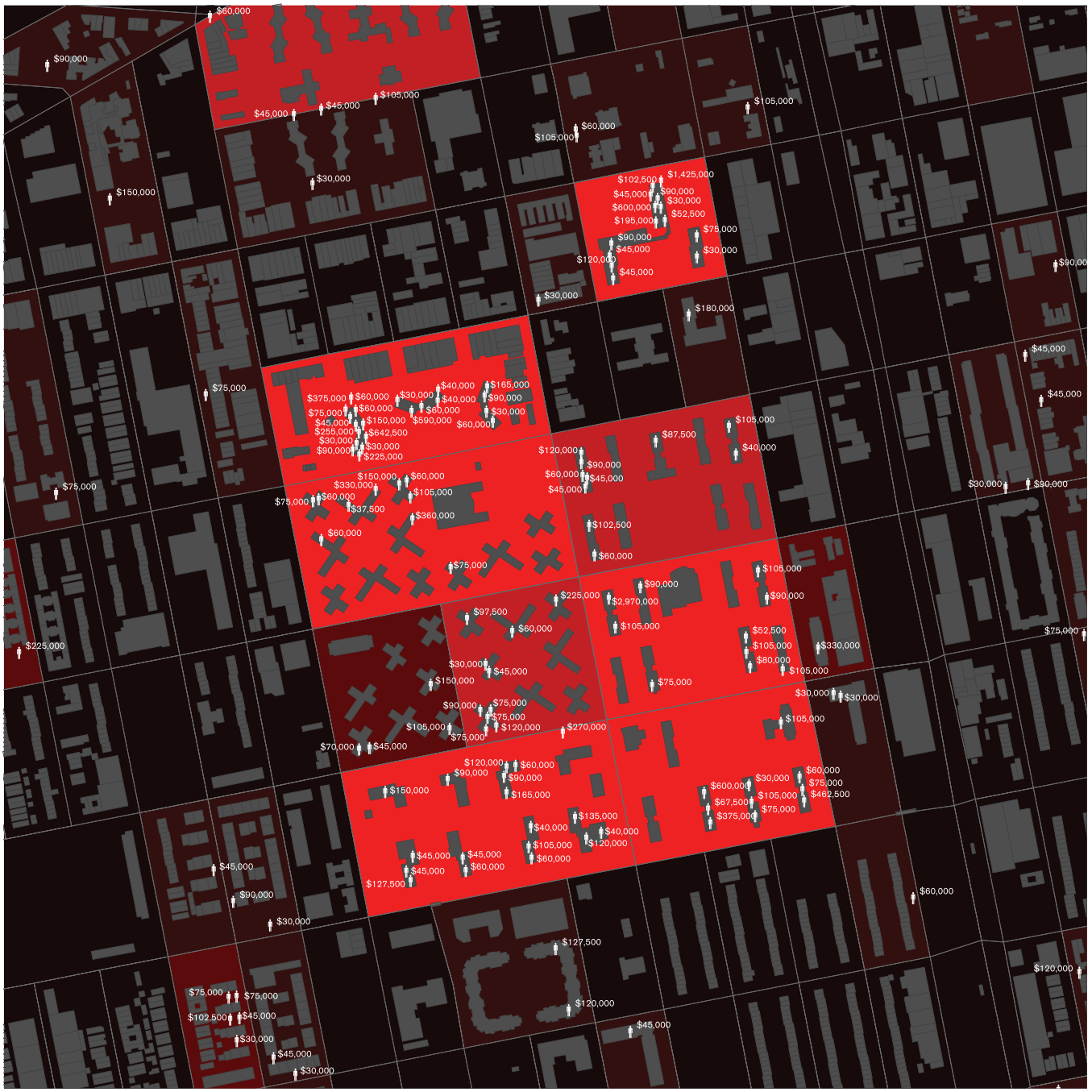
Percent Adults Admitted to Prison, 2003.

BROOKLYN COMMUNITY DISTRICTS	% POPULATION	% POVERTY	% ADMISSIONS
BROOKLYN CD 1	8.51 %	9.08 %	5.37 %
BROOKLYN CD 2	4.03 %	3.58 %	4.64 %
BROOKLYN CD 3	5.89 %	8.10 %	16.51 %
BROOKLYN CD 4	4.24 %	6.34 %	9.34 %
BROOKLYN CD 5	7.04 %	9.30 %	14.45 %
BROOKLYN CD 6	4.22 %	2.60 %	3.08 %
BROOKLYN CD 7	5.02 %	5.03 %	3.82 %
BROOKLYN CD 8	3.78 %	4.15 %	9.46 %
BROOKLYN CD 9	4.26 %	4.14 %	4.43 %
BROOKLYN CD 10	4.96 %	2.79 %	0.91 %
BROOKLYN CD 11	7.05 %	5.54 %	1.35 %
BROOKLYN CD 12	7.39 %	8.54 %	1.32 %
BROOKLYN CD 13	4.23 %	4.94 %	3.41%
BROOKLYN CD 14	6.76 %	8.22 %	3.78 %
BROOKLYN CD 15	6.48 %	4.42 %	1.20 %
BROOKLYN CD 16	3.48 %	5.92 %	8.43 %
BROOKLYN CD 17	6.75 %	5.40 %	5.29 %
BROOKLYN CD 18	7.96 %	3.91 %	3.20 %
BOROUGH TOTAL	100.00 %	100.00 %	100.00 %

Comparisons Expressed as Percent of Borough Total.



Prisoner Migration with Expenditures by Block, 2003.

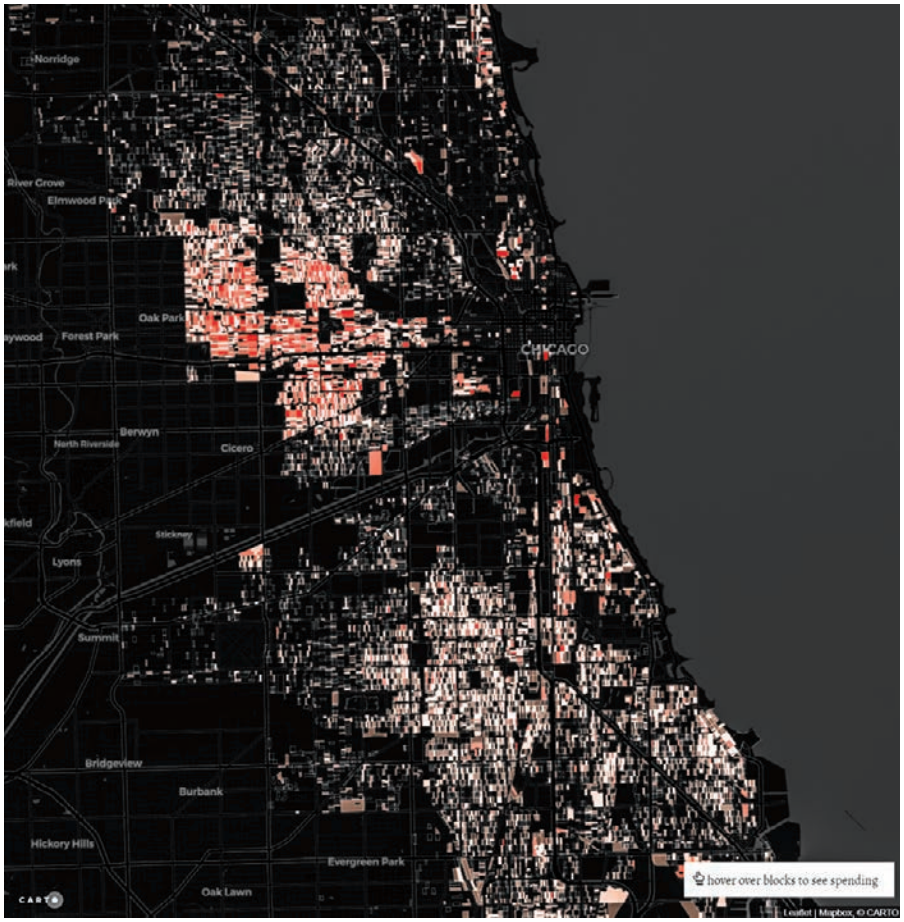


4.13 This map, seen on the wall of the Design and the Elastic Mind Exhibition at MoMA in 2008, shows that more than $17 million was spent on these blocks in Brownsville, NY. Source: Maps created by Eric Cadora, Laura Kurgan, David Reinfurt, and Sarah Williams, courtesy of Spatial Information Design Lab.

The message was delivered using multiple forms of communication. At a 2006 exhibition of the Architecture League in New York City, maps of New York (New York), Wichita (Kansas), New Haven (Connecticut), New Orleans (Louisiana), and Phoenix (Arizona) were presented to the public (figure 4.16). A video seen upon entering the exhibition showed the flow of those incarcerated from their homes in Brooklyn to prisons in upstate New York, which amounted to a mass migration of people. The video asks the public to imagine if just "one million dollars was spent on resettlement rather than imprisonment," what the communities might look like. While the video did not suggest a policy action, it was meant to teach the public about a topic, ask them to rethink it from their own perspective, and ultimately provoke a debate on mass incarceration and its effects on the communities once home to the incarcerated. The Architecture League exhibition helped the project gain recognition of the larger arts and architecture community, and was included in the "Design and the Elastic Mind" exhibition at the Museum of Modern Art in New York in 2008. Now part of MoMA's permanent collection, the maps have been widely exhibited, discussed and disseminated (figures 4.12 and 4.13).

The Million Dollar Blocks visualizations not only supported a public debate through exhibitions, but the work was also published online, in magazines, and other multimedia forums. The simplicity of the maps and image format allowed the message to be easily exchanged through different forms of media. Other groups who sought to make maps of their own city, most notably in Chicago in 2015, imitated the visualization style (figure 4.14).[24] The ability to acquire these images openly online allowed them to be used for all kinds of arguments, including to support congressional funding for the Criminal Justice Reinvestment Act of 2009, which allocated funding to prisoner re-entry programs. The maps brought to light what many criminal justice policy experts already knew, which many in the general public had yet to understand.

The Million Dollar Blocks project could not have been achieved without the collaborative work from our diverse team of criminal justice policy experts, city planners, architects, data scientists, and graphic designers.[25] Acquiring the data came through the relationship that our partners at the Justice Mapping Center had developed during years of work with the Council of State Governments, which works with federal and state corrections departments on a range of policies to reduce costs in state prisons across the



4.14 In 2015, a Chicago organization copied the visual style (red on black) of the Spatial Information Design Lab's 2006 Million Dollar Blocks project. However, not all copies are perfect, and the map has a few flaws, including a missing legend, which makes it hard to know if the data was ethically collected. Source: DataMade, "Chicago's Million Dollar Blocks," Chicago's Million Dollar Blocks, January 7, 2019, https://chicagosmilliondollarblocks.com/.

country. There was much concern about exposing the privacy of the people the data represented, as there should be, and it was our previous work with the Justice Mapping Center that allowed them to trust that we would use the data ethically. Our diverse team, brought together by Laura Kurgan, used its expertise to develop the project, which involved synthesizing data, developing the graphic style, and contextualizing the work with current policies and outcomes of the form of cities. We made our materials available

4.15 Workshops gathered community members, urban planners, criminal justice policy experts, and data scientists to ask how they might use just $1 million to change Brownsville in Brooklyn, where an overwhelming majority of prisoners come from. Source: Maps created by Eric Cadora, Laura Kurgan, David Reinfurt, and Sarah Williams, courtesy of Spatial Information Design Lab.

4.16 Exhibition at the Architecture League. Source: Maps created by Eric Cadora, Laura Kurgan, David Reinfurt, and Sarah Williams, courtesy of Spatial Information Design Lab.

in several formats, helping the work spread further as it tapped into multiple networks and delivering the message of the Million Dollar Blocks project to broad groups of people.

The team engaged in extensive discussion on the ethics of the presentation of the data—which included a discussion of potential biases. The maps were meant to persuade, so we were well aware of our inherent bias. Therefore, we developed numerous versions of the maps making sure everything—from the way we binned the colors in the legend to how we normalized the data—was presented responsibly. We considered it essential to have criminal justice policy experts evaluate the ethics of the maps, as we did not want to further marginalize the populations represented in the maps.

More than a decade has passed since our team developed the maps, and I continue to learn about the varied ways the message of the project impacted others. In July 2019, I sat on a panel with Ifeoma Ebo, a Design Advisor for the Mayor's Office of Criminal Justice in New York City; she spoke about work she is doing to help people returning from prison, and she mentioned Million Dollar Blocks as an influence. Million Dollar Blocks was an inspiration not only to the government of New York City but also to criminal justice advocates as far away as Australia, where the government used the research and maps as evidence for justice reinvestment policies in 2018.[26] The messages of maps do have the ability to travel far.

Interactive Websites Help to Create User-Centered Data Narratives

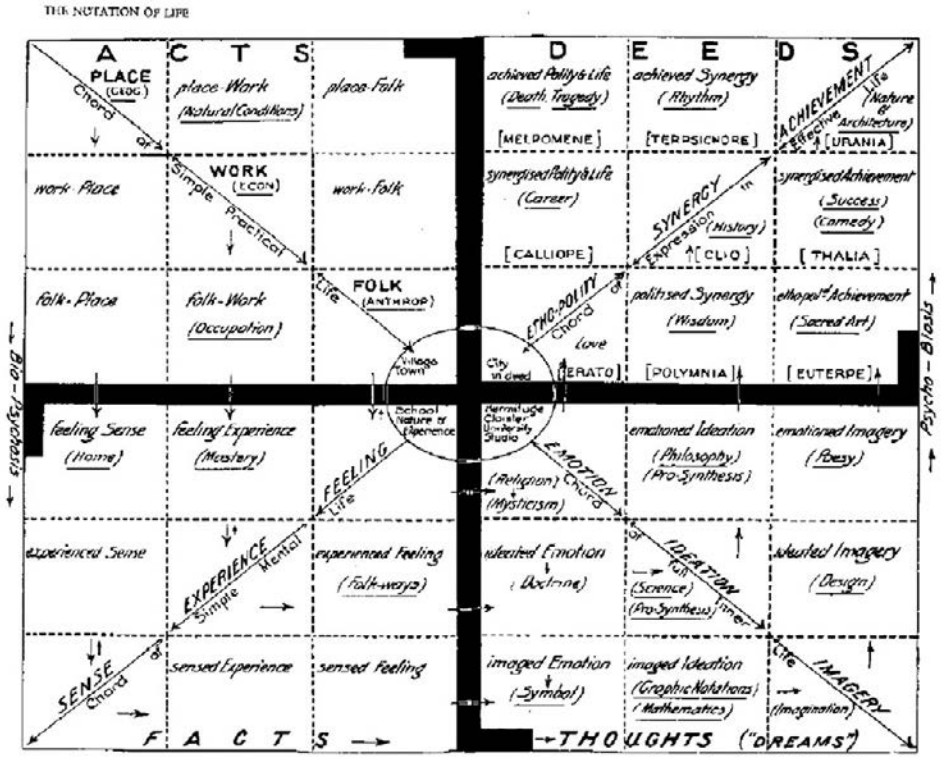
Much of the criticism revolving around the power dynamics and bias associated with visualizations occurs because it is hard to communicate the complexity of the world through data visualizations that can only provide one vantage point or interpretation. All data has multiple interpretations. Some early innovators in the fields of data for society struggled with this as well. Perhaps the most notable among them was Patrick Geddes (1854–1932), a social statistician of the Scottish tradition and a contemporary of Charles Booth, maker of the famous Booth maps mentioned in chapter 1. Geddes believed surveys were essential to town and regional planning, but he felt limited by the one-dimensional nature of many surveys of his time and encouraged the development of cadastral map sets that included "geography, geology, climate, economic life, and social institutions."[27] Geddes struggled with the limitations of maps as a representational tool; he believed they

could not provide the real interplay between the various dynamics that make cities thrive.[28] Contemporary data scientists have more ways to present data: web maps and interactive visualizations, now commonplace, allow us to explore these relationships and find our own meaning in the data. One can imagine how Geddes might have enthusiastically deployed these new digital technologies to illustrate his beliefs on the interface of nature, built form, and its inhabitants.

Geddes's calls, in writings and lectures, for new representational strategies often fell on deaf ears, mostly because his arguments were extremely difficult to follow—and he was said to have a difficult and argumentative personality.[29] Yet, at the heart of his position was the belief that to make planning decisions we must truly understand the people, the geography, and how the two interact, all the while involving the public in the practice of observation. These ideas have had a lasting impact on regional town planning and have influenced many regional planners (including Lewis Mumford) in addition to the methods laid out in these pages.[30]

Geddes is one of the first to bring out the limitations of data analytics, namely the problem of trying to quantify the qualitative nature of cities. While he himself advocated for the development of data, he investigated alternative ways of communicating the insights found within through data visualization. His most notable method of melding the psychological, social, and physical elements that make a city successful is his diagram "The Notation of Life" (figure 4.17), which charts how the relationships between a city's physical form and its citizens' social interactions can result in a flourishing civic life.[31]

Geddes's struggle with the one-dimensional presentation of data provided by maps is similar to that of today's Critical Cartographers, who believe we must interrogate the reasons maps are developed in order to expose the power dynamics that might not otherwise be apparent on a visual reading. Earlier in this chapter I mentioned Jeremy W. Crampton, who argues in his article "Maps as Social Construction: Power, Communication, and Visualization" that one way to account for the power dynamics and inherent bias in the representation of data is to develop ways in which map readers can explore the data and understand it from their own perspective. Engaging data in this way has been made easier with the development of interactive websites that facilitate data exploration. Crampton also believes cartographers should be more sensitive to how the public might interpret the information in maps



4.17 Patrick Geddes's "The Notation of Life." Source: Geddes, *Cities in Evolution* (London: William and Norgate Limited, 1949).

as misleading or marginalizing. He asks mapmakers to critique their maps through the eyes of the public.

Local Lotto: A Web Tool for Data Literacy

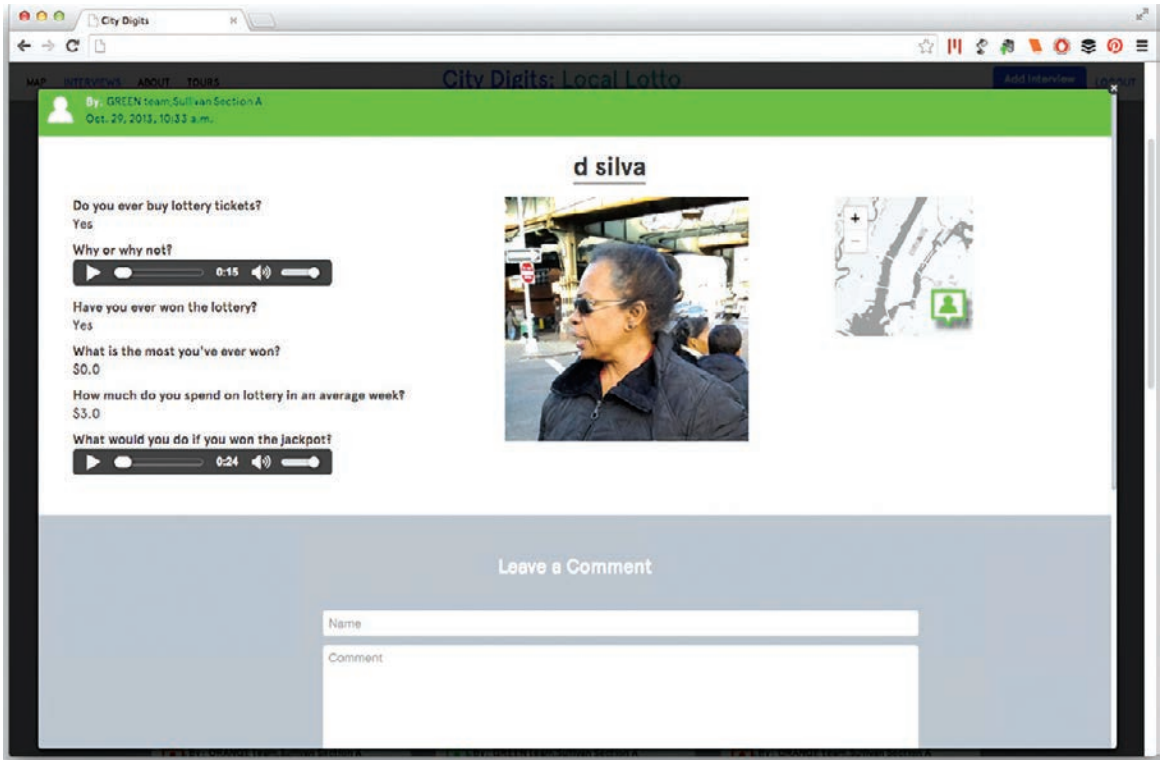
Narratives, images, and interviews are also data, and figuring out how to bring these diverse data sets together is important for generating evidence to help change policy. This is exactly what a group of high school youth did in 2014 when they participated in Local Lotto, a module of a math curriculum called City Digits, to create arguments for whether the lottery provided a benefit to their community. The City Digits curriculum, a collaboration with my research lab (Civic Data Design Lab), CUNY's Brooklyn College, and the Center for Urban Pedagogy (CUP), is developed on a web-based interface that helps high school students build data literacy by allowing them to collect, explore, and form opinions about social justice topics they observe

in their neighborhood every day. The Local Lotto module gives students the opportunity to investigate quantitative and qualitative data about the state lottery using interactive maps and participatory data collection. Beyond supporting youth to form opinions about a local and socially relevant topic, Local Lotto's interactive format led to new forms of equitable classroom participation and learning (see figures 4.18–4.24).

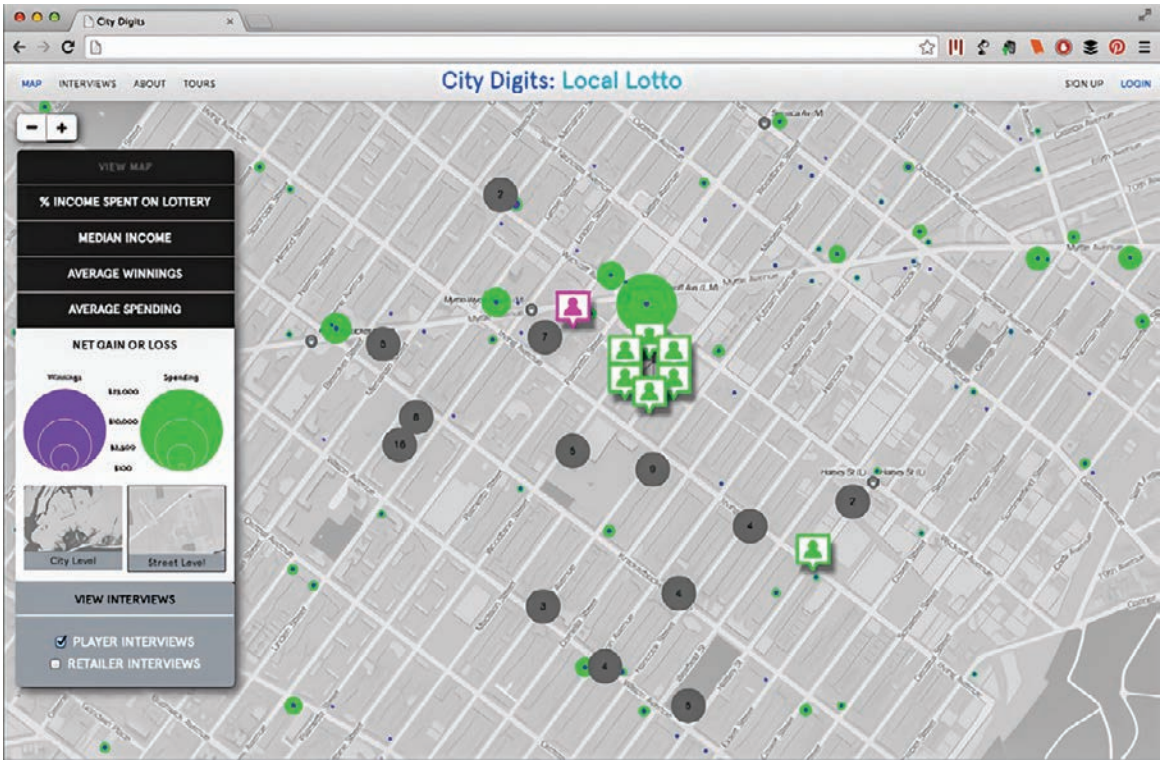
Local Lotto comprises four sections that use an interactive web tool. First, students learn how to calculate the probability of winning a jackpot lottery game. Second, they conduct and collect interviews of lottery players and retailers in their neighborhood using the Local Lotto tool on mobile tablets (figure 4.18). Third, students analyze citywide and local level lottery data obtained from the New York State Lottery Commission and public data from the 2010 Census, using an interactive map. Fourth, students synthesize qualitative interview data with quantitative map data to formulate their own opinions about the lottery's social impact. Using the web-tool, students create multimedia narratives called "tours" to teach others about what they learned. Developing the tours allows students to synthesize their data explorations and form opinions, which is an important component of data literacy (figure 4.24).

In Local Lotto, thematic maps highlight the social dynamics of the lottery by visually communicating quantitative data overlaid on a map of the city by neighborhood. Ultimately, the maps show a pattern of increased lottery spending, relative to income, in low-income neighborhoods. Photographs of the local streetscapes and interviews with pedestrians allow for an interpretation of this pattern from a first-person perspective (figures 4.21

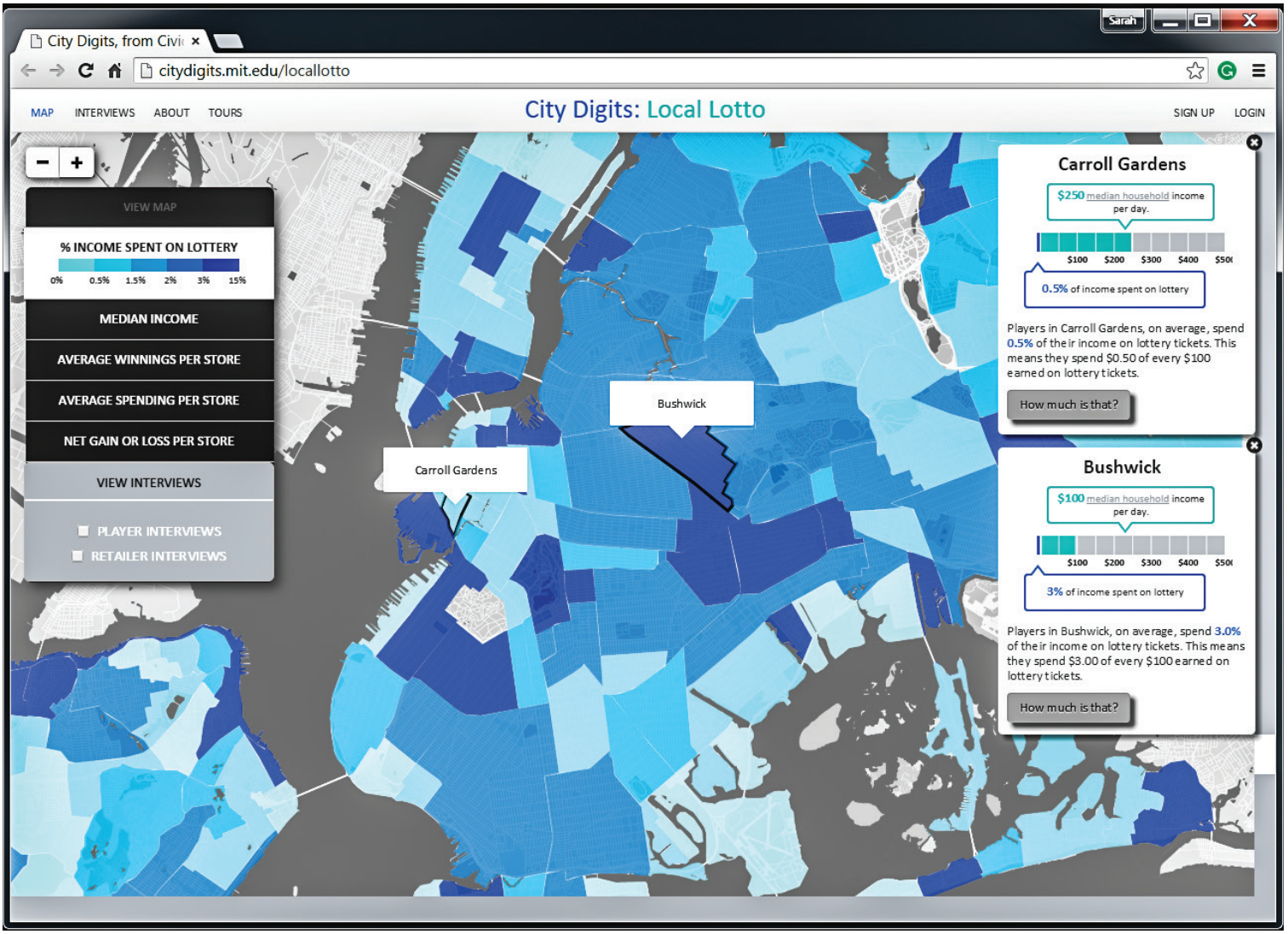
4.18 Students who participated in the Local Lotto curriculum conduct interviews in their neighborhood. Source: Laurie Rubel.



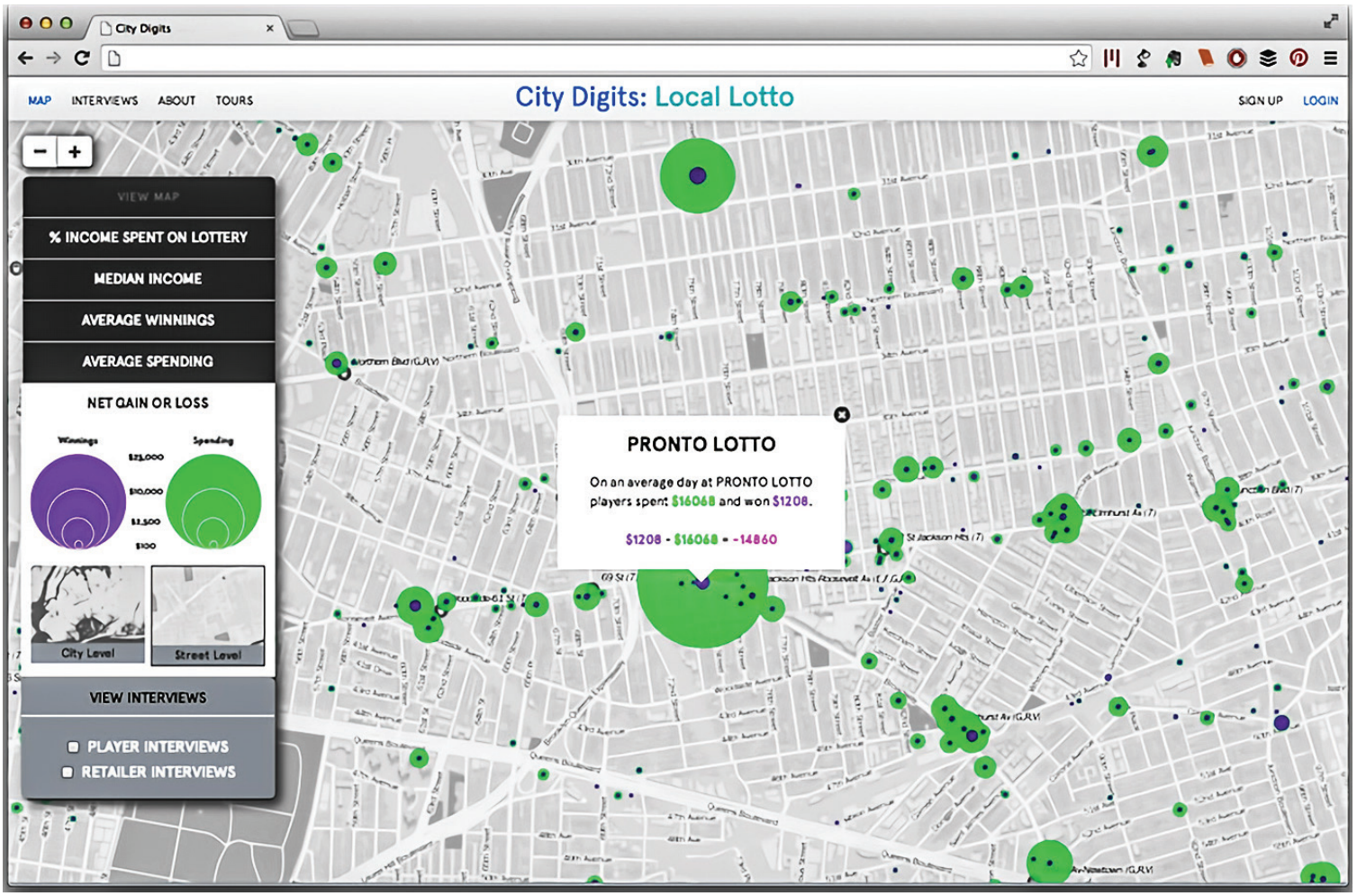
4.19 As part of the Local Lotto curriculum, students interviewed people in their neighborhood who played the lottery. The interviews went directly to the map as a voice recording of answers to questions that the students developed. Source: Screen shot courtesy of Sarah Williams and Laurie Rubel, City Digits, http://citydigits.mit.edu/.



4.20 Interview icons sit on top of a map of lottery data. Source: Screen shot courtesy of Sarah Williams and Laurie Rubel, City Digits, http://citydigits.mit.edu/.



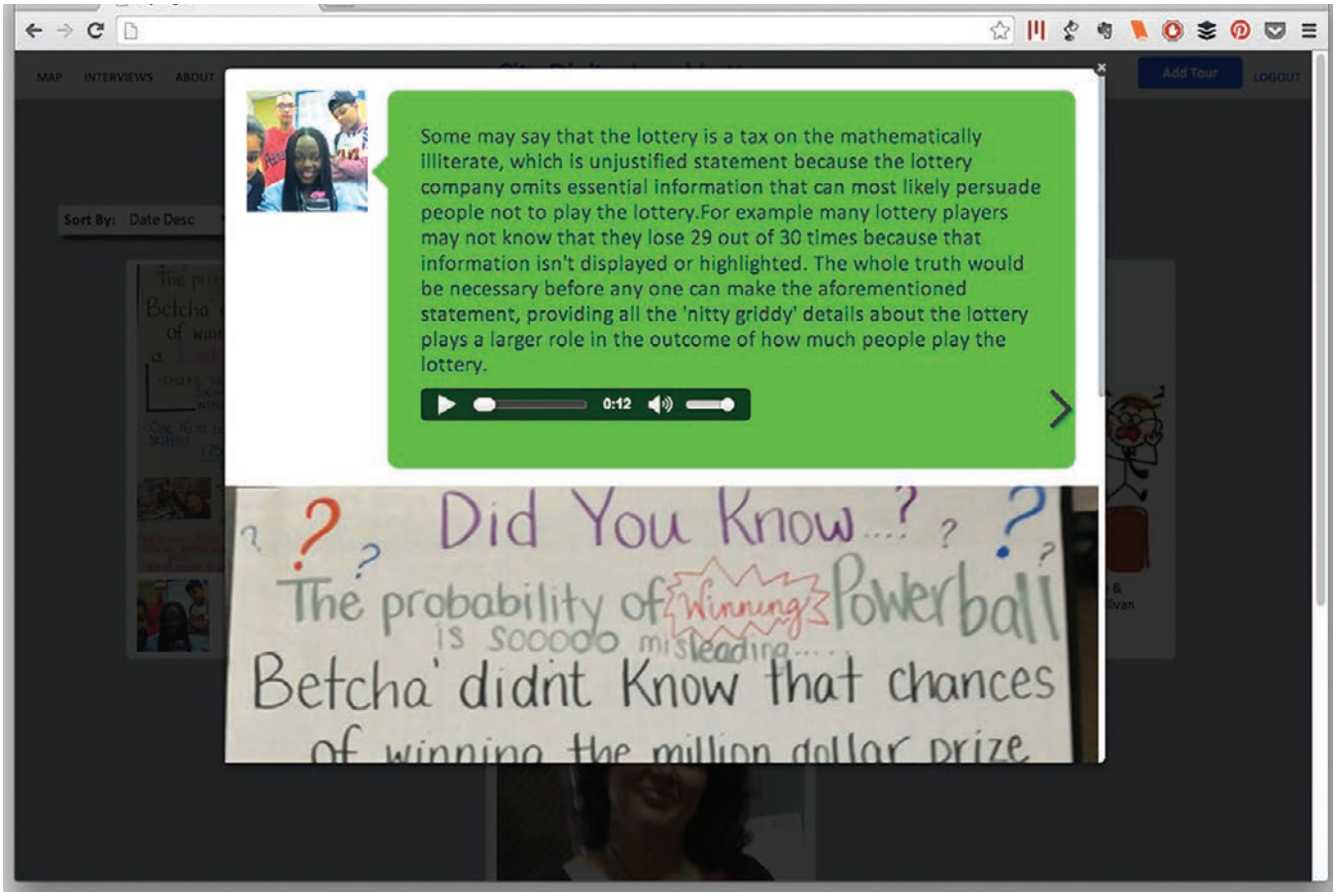
4.21 This Local Lotto map shows the percentage of income spent on lottery tickets. The darker blue represents areas where a greater percentage of median household income is spent on lottery tickets. Source: Screen shot courtesy of Sarah Williams and Laurie Rubel, City Digits, http://citydigits.mit.edu/.



4.22 Net gain or loss map, zoomed into the street level. The green circles represent the amount of money spent on lottery tickets at individual retail stores. The purple circles represent the amount of money won from lottery tickets at the same stores. Source: Screen shot courtesy of Sarah Williams and Laurie Rubel, City Digits, http://citydigits.mit.edu/.

4.23 Students taking part in the Local Lotto curriculum work together to construct arguments using data as evidence. Source: Laurie Rubel, City Digits.

4.24 This Local Lotto student tour explains that lottery marketing represents an "omission of the truth." Source: Screen shot courtesy of Sarah Williams and Laurie Rubel, City Digits, http://citydigits.mit.edu/.



through 4.24). Combining geographic analysis with images and interviews contextualizes data and allows for the investigation of a civic issue from multiple vantage points.

Local Lotto focuses on teaching data literacy by combining quantitative data (in the form of thematic maps) with qualitative data (in the form of student-collected interviews) and photographs. Blending quantitative and qualitative investigation enables a more nuanced analysis of the lottery, and the use of maps helps reveal relationships between geography and social issues. One student group argued that the lottery was a "scam because it targets low-income areas."[32] Using a map to make the argument they combined the information with interviews they had recorded, in which local deli owners mentioned they thought it was sad that customers with so little money spent it on the lottery.

Inherent in teaching data literacy is the ability to generate arguments about quantitative and qualitative data.[33] In Local Lotto, these arguments tackle whether the lottery is "good" or "bad" for communities. The web tool facilitates the development of argument by allowing students to explore the data on a map. The discussion that ensues while students construct their arguments helps to teach them the policy issues behind the lottery and create a dialogue that is important for community engagement.[34]

The interactive and participatory nature of the curriculum and tool allowed youth to deeply explore data they could connect with. Students indicated that they worked harder in the Local Lotto class than they did in regular mathematics classes because the content was both challenging and relevant to their own lives. Students who were learning English at the same time, and who are typically challenged to participate in class, were able to take leadership roles in the data collection. Some students reflected on the personal significance of their learning during Local Lotto and indicated that they shared their findings at home with family members who regularly buy lottery tickets. Local Lotto taught the youth math as well as how to debate a topic important to society. It also taught students how to take control over narratives around data by allowing them to create their own stories.

Data Stories Create Evidence for Public Debate

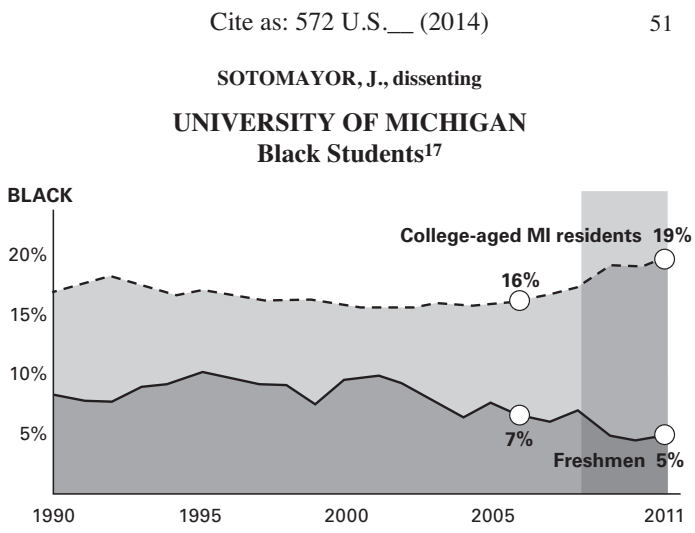
Newspapers and other media have long been used as forums for public debate, and many news agencies have begun to use data visualizations as a way to

communicate policy issues. Online media's ability to develop interactive "data stories" allows users to explore data, ask their own questions, and generate new conversations. These visualizations do not provide single story lines; rather they allow readers to find their own. The resulting narratives help create a debate on civic topics.[35]

One notable case in which a *New York Times* interactive map generated public debate was an infographic titled "How Minorities Have Fared in States with Affirmative Action Bans," to the point that the graphic appeared in documents used in the Supreme Court. The graphic, developed in April 2014, allowed readers to explore the percentage of Hispanic and black students accepted to state universities in Washington, Florida, Texas, California, and Michigan before and after affirmative action bans were established. At the time the graphic was published, the Supreme Court was set to hear the case of *Schuette v. Coalition to Defend Affirmative Action* (2014), which would decide whether to uphold the affirmative action ban that had been in place in Michigan since 2006. The *Times* generated the visualization so readers could better understand the questions put before the Supreme Court. Overall, the charts and accompanying graphics showed decreases in minority enrollment after affirmative action bans were set in place.[36]

When Supreme Court Justice Sonia Sotomayor used the graphic as part of her official dissent to the ruling to uphold the affirmative action ban, the work became part of a historic debate. She used a screenshot of the *New York Times* visualization to support her position that the removal of affirmative action does, in fact, affect minority enrollment in college (figure 4.25). She stated: "The proportion of black students among those obtaining bachelor's degrees was 4.4 percent, the lowest since 1991."[37] Using the graphs as evidence, she went on to say, "At UCLA, for example, the proportion of Hispanic freshman among those enrolled declined from 23 percent in 1995 to 17 percent in 2011, even though the proportion of Hispanic college-aged persons in California increased from 41 percent to 49 percent during that same period."[38] Sotomayor has always been a supporter of affirmative action, so the graphic most likely did not inform her decision, but she hoped to persuade others by illustrating through data the relationship between policies and minority enrollment.

The *New York Times* received some criticism that these graphics were biased toward lifting affirmative action bans. One article in the *Independent Voter Network* claimed, "All of these numbers have left off something



4.25 Screenshot from Justice Sonia Sotomayor's dissent in the Supreme Court case to uphold a Michigan ban on affirmative action. It is a direct copy of a map she saw on the topic in the *New York Times*, showing how newspapers can help to create public debate. Source: Sonia Sotomayor, *Schuette v. Coalition to Defend Affirmative Action*, No. 572 U.S. 2014.

very important; the students that leave those states to attend a university in another state and the students from other states that go to those particular universities. Should they be used to help show if a university is measuring up in minority enrollment to the state's total minority population?"[39] Others found the arguments misleading because they believed the data should have compared college-ready minority students rather than college-age students.[40] Still other articles point out that the *Times* did not include Asians or other minorities in their definition, and, therefore, the story was misleading.[41] Those who supported the Supreme Court's decision often reused the same data visualizations to argue their position. The graphics' appearance in the *New York Times* caused them to be reappropriated and used as evidence for different stories. In short, the graphics generated a debate, which is key to public engagement.

"Graphic storytelling has become an important part of journalism," says Jeremy White, a graphics editor for the *New York Times* who creatively incorporates pictures, video, narratives, graphs, and maps into the stories he develops for the *Times*. The web helps to bring these multiple forms of communication together, allowing users to experience both qualitative and qualitative narratives. The *Guardian*, one of the United Kingdom's main newspapers, does a great job of this kind of storytelling as well, and has developed numerous visualizations that have generated public debate, even here in the United States. One such visualization was "The Counted" (2017),

which has collected data on deaths by members of the police force across the United States. *Guardian* readers can slice and dice the data in various ways, finding new insights and reading the narratives from the news stories about the incidents (figure 4.26). What's interesting about "The Counted" is that it did so much more than present the startling story of police killings, but that the data, systematically collected by the *Guardian*, served to change US policies. The *Guardian* mined data from websites, looking for incidents of assaults, and found many more than were reported by federal and state agencies across the country—a marked underreporting. The problem is indeed so widespread that the true cause of death often doesn't make it to the National Vital Statistics System (NVSS). A 2017 Harvard study showed that NVSS was undercounting 55.2 percent of these deaths and that the errors were higher in low-income neighborhoods.[42] The *Guardian* compared its database to those of the Bureau of Justice Statistics (BJS) and even FBI reports, and it reported again on a mismatch. The press about the topic caused all of these agencies to change how they reported data by developing visualizations.[43]

4.26 Screenshot of the "The Counted: Tracking People Killed by Police in the United States." Source: *Guardian*, 2016, https://www.theguardian.com/us-news/series/counted-us-police-killings.

Stop-and-Frisk Brought to an End Using Data

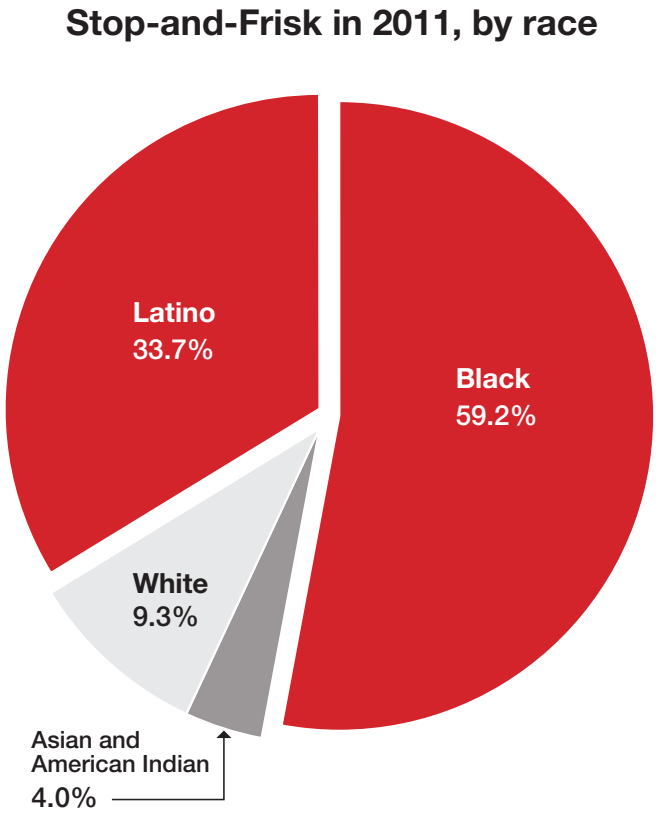
Sharing data has the power to create the public outcry needed to eliminate discriminatory government and law enforcement policies. Under President Obama's open data policies, local and federal governments made available previously inaccessible data, and the public used this data to counter government narratives and change policy.[44] A great example is the New York Police Department's (NYPD) stop-and-frisk data, or Stop, Question, Frisk (SQF), as it was referred to by New York City Mayor Michael Bloomberg, who enacted the policy. This policy allowed police officers to question and search individuals for anything the officers considered to be "suspicious activity"—a category interpreted broadly and with high subjectivity.[45] The policy came into contention when data analysis of UF-250 forms, which were used by the police to report these stops, showed that only 20 percent of stops were warranted. It should be noted that a later review of the data by a federal judge in 2013 found the figure conservative—in reality, most stops were not warranted, showing how the SQF policy was being abused by the NYPD to unduly search members of the public.[46]

Perhaps what is more interesting about the stop-and-frisk story was not the evidence presented at the court case, but rather the fact that evidence was obtainable at all. In 2008 the Center for Constitutional Rights (CCR) performed an analysis of stop-and-frisk data from 2003 to 2007 that was available through new open data policies. They found that stops increased 200 percent in just those four years, and that of those stops, 88 percent did not lead to crimes. They also found that 85 percent of the people stopped were black or Latino, news that gave rise to concerns about racial profiling. In conversations with policing advocacy groups the CCR decided they needed to take action. In 2008 the CCR filed a class action lawsuit, *Floyd, et al. v. New York City, et al.*, arguing that the police were using racial profiling tactics when determining which people on the street to stop.[47]

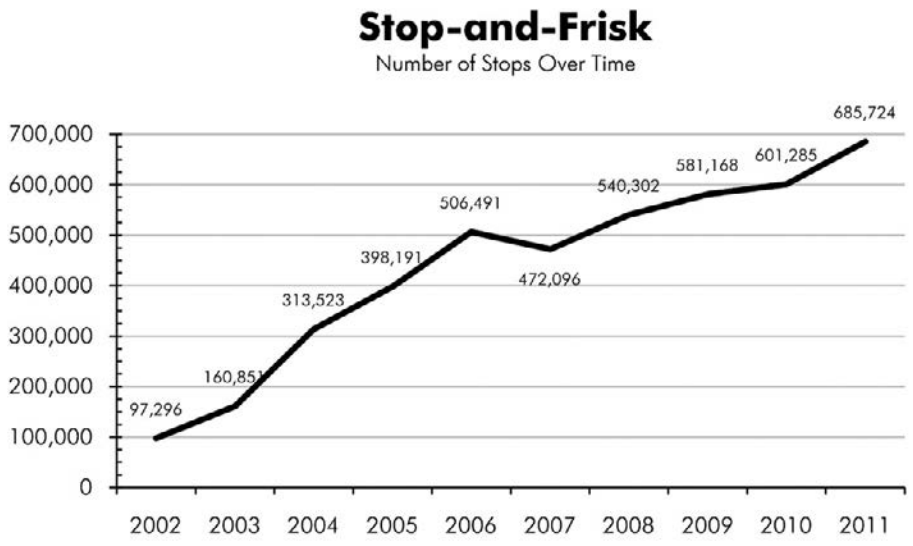
CCR's mission is to use litigation to fight for social justice, and as they started their work on stop-and-frisk they engaged numerous organizations already working toward police reform. These organizations ultimately recognized their shared interest and created the Communities United for Police Reform (CPR) in 2011 to coordinate as one unified organization to advocate for the process.[48] CPR worked to prove the legitimacy of the case, but also to educate the public using the skill sets of its diverse members, each using their individual advocacy techniques.

The New York Civil Liberties Union (NYCLU) performed additional analysis on 2011 NYPD data, again showing the police were stopping higher proportions of minorities. A report they released in May 2011 found that of the 684,330 people stopped in 2011, only 12 percent were arrested or received summonses.[49] This helped create a counter-narrative to the police department's argument that the policy helped prevent crime. The report, which used visualizations to explain the findings, was widely discussed in the media, including mentions in the *New York Times*, *Forbes*,[50] and the *Wall Street Journal*, among others. These press campaigns helped to get the topic on the agenda of the next mayoral campaign and it was heavily discussed during debates. Bill de Blasio, who eventually became New York City's mayor in 2013, campaigned to do away with the policy.

Beyond the initial visualizations created by the New York Civil Liberties Union, media outlets across New York City used the same stop-and-frisk

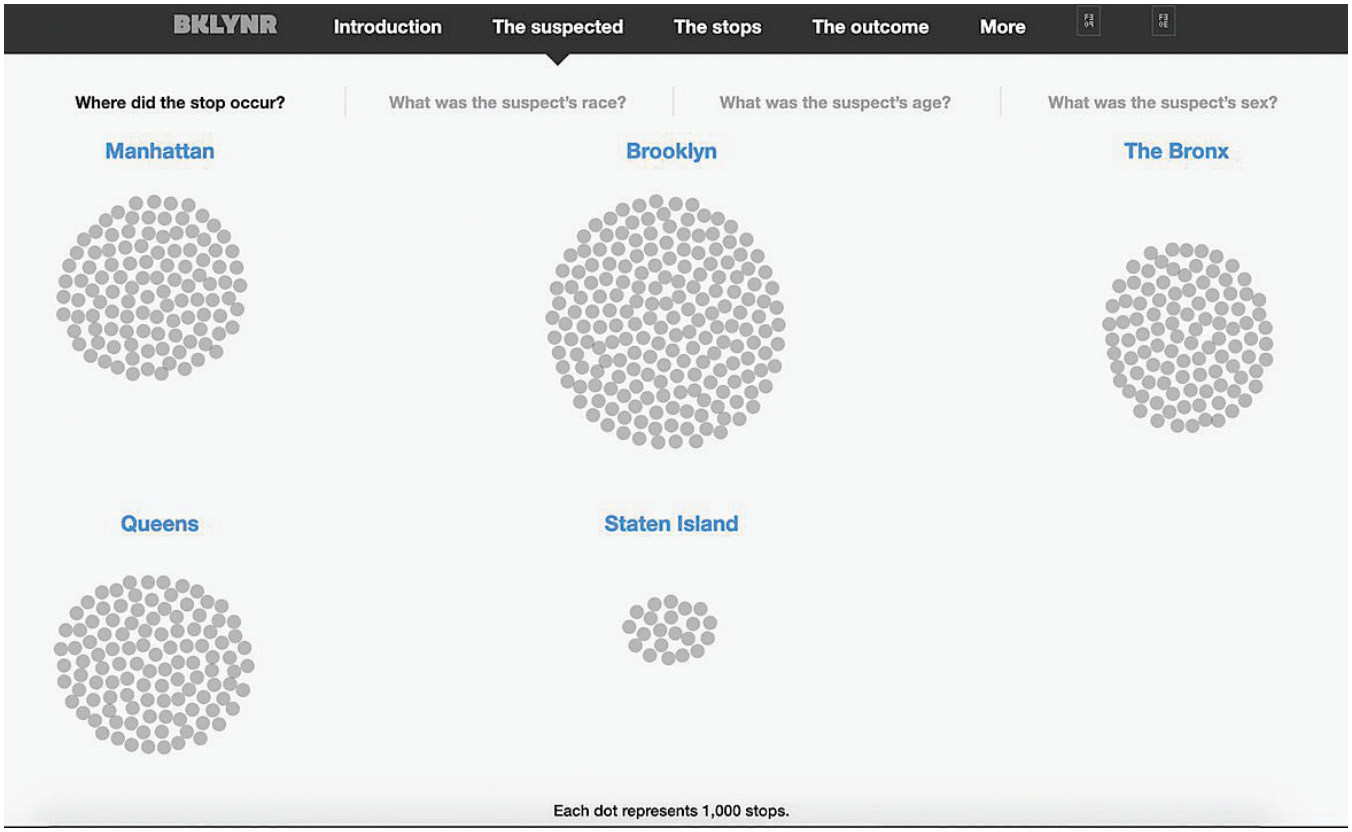


4.27 This pie chart from the NYCLU May 2011 report on stop-and-frisk shows that the majority of people stopped were Latino or black.



4.28 This graph from a May 2011 NYCLU report shows the rise of stop-and-frisk practices.

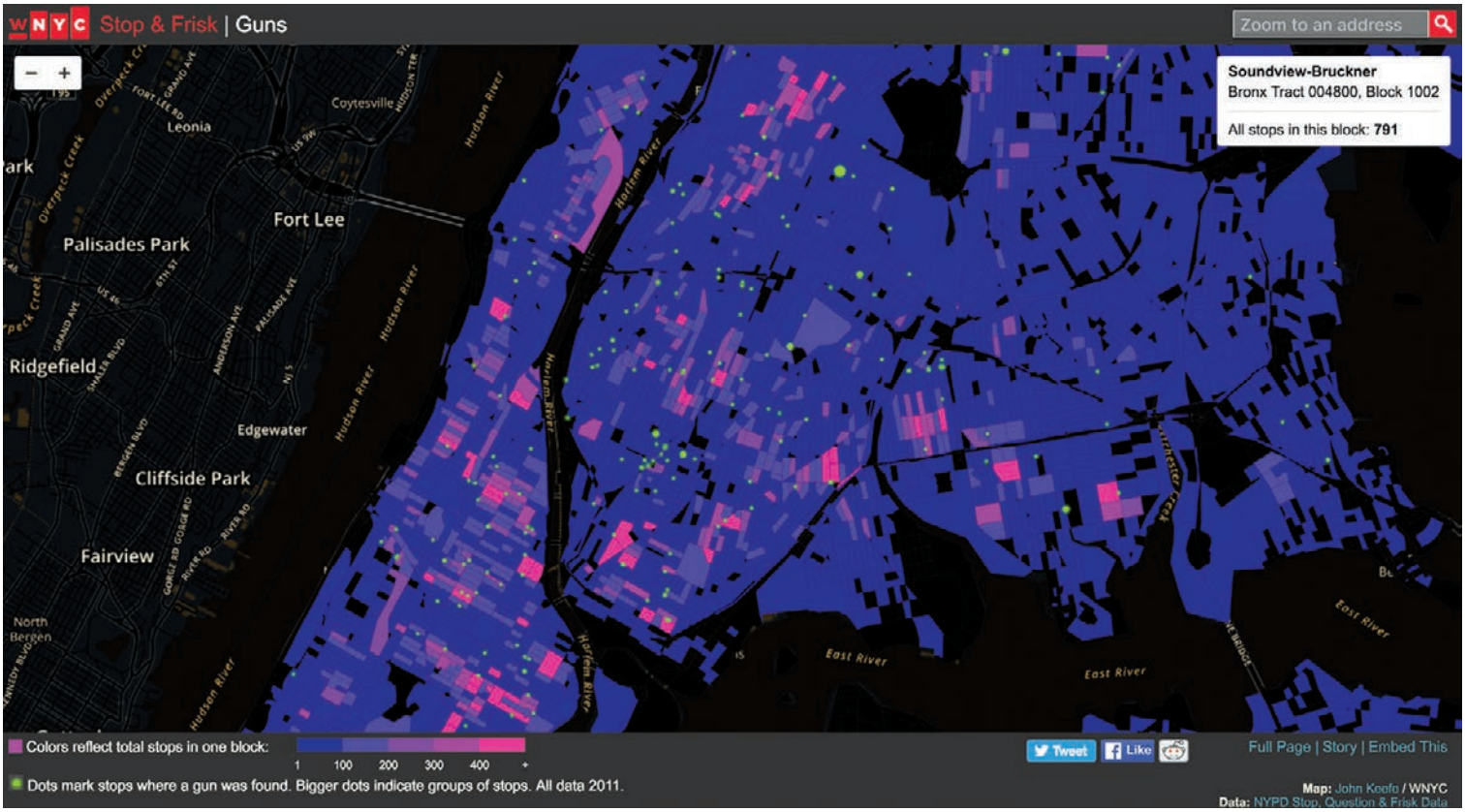
data to help illustrate this unjust police practice. This couldn't have happened were the data not openly available. An independent magazine on Brooklyn issues, *BKLYR*, created a visualization that allowed users to break down the stop-and-frisk data using multiple variables online (figure 4.29).[51] The data visualization was widely circulated among news outlets including a feature in *Wired* magazine; it was retweeted, spreading it across the web. WNYC, New York's flagship public radio station, created the most mentioned visualization, a map (figure 4.30) on which the hot pink blocks represent over 400 stop-and-frisks, or an average of more than one a day. The visualization shows how some minority communities are targets for the policy.[52] It should be noted that the WNYC visualization has also been criticized for its representation style because it has led some map viewers to believe that there are more crimes and guns in these neighborhoods. It was not made clear to the viewers that people in those neighborhoods were stopped more frequently (and often unjustly so), creating a bias in how the data was represented through the map. It is therefore important to be careful in how you construct the narrative around data visualizations. There were also more experimental visualizations, such as the one by artist Catalina Cortazar; she created something akin to a sand



4.29 BKLYNR blog's stop-and-frisk data visualization, in which each dot represents a proportion of the population so users can easily and visually interact with the data. Source: "All the Stops," BKLYNR, https://www.bklynr.com/all-the-stops/.

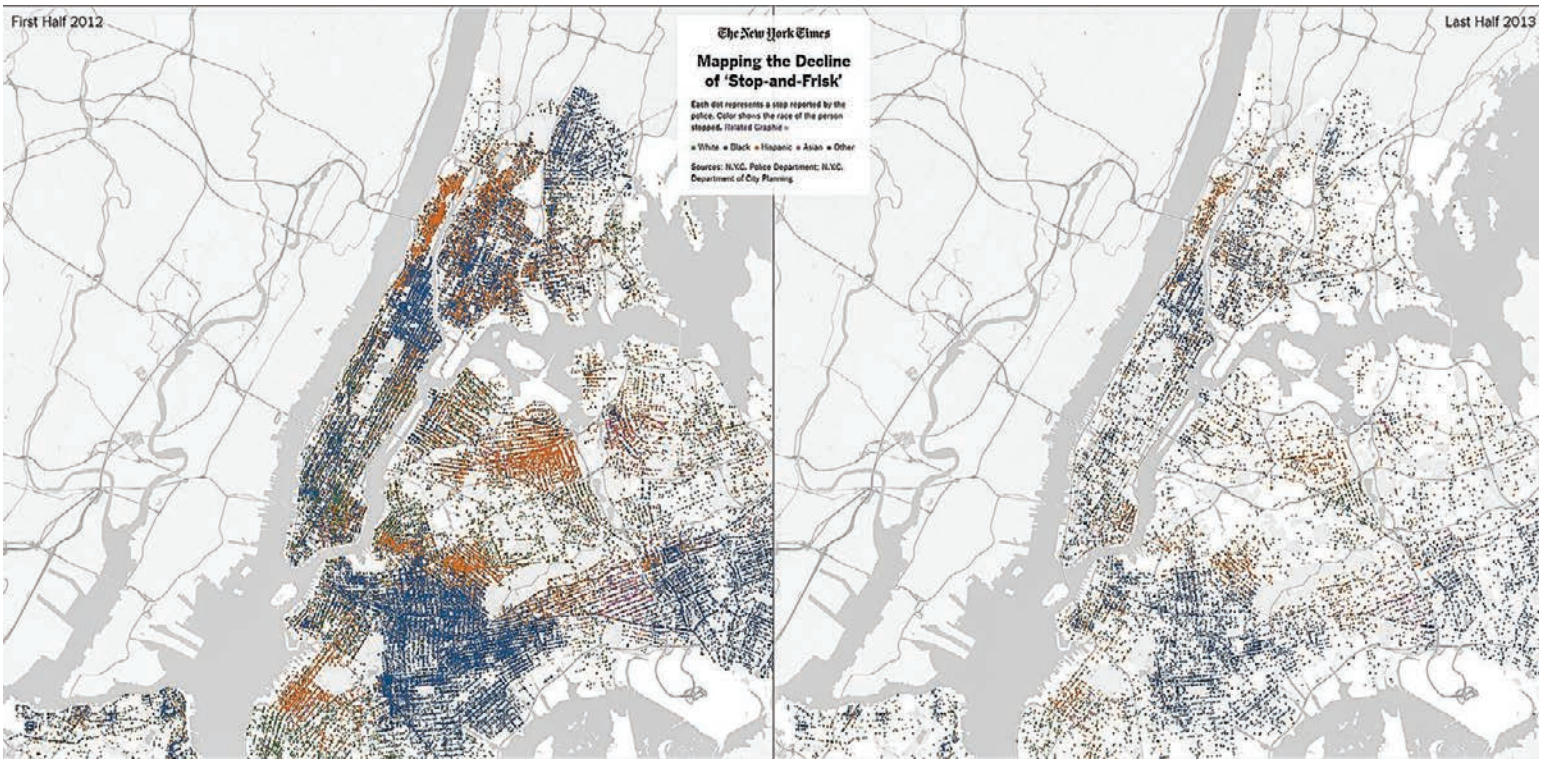
clock, and when she turned the sculpture upside down, the sand that couldn't fall through represented the proportion of the population of whites, blacks, and Latinos who were stopped in New York City.[53]

Ultimately CCR filed a federal court case against the policy, and the trial lasted nearly nine weeks. In May 2013 the court ruled the policy unconstitutional under the Fourth and Fourteenth Amendments.[54] "[The City has] received both actual and constructive notice since at least 1999 of widespread Fourth Amendment violations occurring as a result of the NYPD's stop and frisk practices," the decision stated. "Despite this notice,



4.30 WNYC produced a map of the openly available data, which was widely spread across the internet. This map was criticized for trying to make a correlation between Stop-and-Frisk and locations where people had illegally obtained guns. The problem is that Stop-and-Frisk tickets were disproportionately issued in certain neighborhoods specifically because the tactic was overused in minority neighborhoods. Therefore, the data is biased in a way that makes it appear as though minority neighborhoods have more guns. Source: John Keefe, "Stop & Frisk | Guns," WNYC, 2013, https://project.wnyc.org/stop-frisk-guns/.

they deliberately maintained and even escalated policies and practices that predictably resulted in even more widespread Fourth Amendment violations. . . . The NYPD has repeatedly turned a blind eye to clear evidence of unconstitutional stops and frisks." The document further demonstrated that evidence clearly pointed toward systematic racist practices.[55] The continuation of SQF was found illegal and the New York City Police Department stopped the practice. In 2014 the *New York Times* created a feature that showed what removing stop-and-frisk looked like, neighborhood to neighborhood, illustrating a huge reduction in stops (figure 4.31). The policy tarnished the



4.31 The *New York Times* data visualization that compared stop-and-frisk after a 2013 federal court case declared the policy unconstitutional. The colors represent the race of people who were stopped (green for whites, blue for blacks, orange for Hispanics, pink for Asians, and brown for others). Source: Mike Bostock and Ford Fessenden, "'Stop-and-Frisk' Is All but Gone from New York," *New York Times*, September 19, 2014, https://www.nytimes.com/interactive/2014/09/19/nyregion/stop-and-frisk-is-all-but-gone-from-new-york.html.

reputation of Mike Bloomberg during the 2020 presidential campaign, as the public found his initial support of the policy hard to reconcile.

The ways in which data was involved in removing New York City's stop-and-frisk policy shows how data developed for power and control can be transformed into something that becomes a tool to overturn unethical practices. It was not merely sharing the data but also the narrative developed around the project that allowed for this reversal. Built collaboratively with policy experts and the community, and delivered to their vast networks, it captured the attention of the mainstream press and media. In turn, the media helped expand the number of people who were aware of the policy and its harms, making it a topic of public debate. Represented in this example are the Data Action methods discussed throughout the book. The stop-and-frisk case shows how anyone can creatively find openly available data, collaborate with experts to tell the data's story, and communicate the result to the public through graphics, ultimately generating societal change.

Sharing Data Creates Transparency, Public Participation, and Collaboration

President Obama once stated that open government seeks "transparency, public participation, and collaboration."[56] Data for policy action needs to abide by those same principles to make ethical uses of data strategies to maximize policy change. Sharing data creates transparency, and this allows those with whom you are sharing data to build trusting relationships that can further mutual policy initiatives. This was true for the Digital Matatus project—the more we shared data with our partners, the more trust we created from and for our mutual work.

Transparency also means exposing data and all the biases within it. This means that through analysis of the data we can illustrate underlying systems of control. Both the prisoner intake data of the Million Dollar Blocks project and the stop-and-frisk story show how you can take a data set used for power and control, and transform it for use by the people whom it represents to exert their own power. In both of these cases the over-representation of communities of color in the data suggests racial bias in policing. In the case of the Million Dollar Blocks, many of the neighborhoods lack the resources to pull themselves out of a cycle of incarceration—jail is the policy answer. In the case of stop-and-frisk, the neighborhoods and the people who live in them

are seen as suspicious, which heightens the risk of over-policing and targeting communities of color.

Sharing data through visual interpretation might make underlying policies more transparent, but all data visualizations hold the bias of the creator. This is why public participation is essential for using data visualizations to create policy action. By asking the people who have a stake in the visualization to scrutinize the results, we are made more aware of the potential biases in the data visualization.

Building data visualizations with collaborators helps us identify our unconscious biases, creates trust in the message the visualization is trying to communicate, and, most importantly, disseminates that message to a wide audience, giving our work with data a greater overall impact. Such collaborative effort is the special sauce for most of the projects described in this book. In Nairobi the co-creation of the maps helped to ensure the graphic language was accessible and also built trust among stakeholders, so much so the government trusted the map enough to make it the official map of the city. With Million Dollar Blocks, by collaborating with policy experts we were able to create consumable media messages about the topic. Our collaborative efforts helped tap into a network of multiple advocacy groups, which helped to get the story out to broad networks.

Sharing data also gives it a life beyond the purposes it was originally intended. By putting their data in the public domain, for example, the NYPD allowed it to be used for other initiatives—specifically understanding trends in racial profiling. This is true in Nairobi, too, where freely sharing the data meant the local technology community could develop its own technologies with the data.When the New York State Lottery commission shared data with us to teach youth about the policies regarding the lottery, they transformed something used for management into a learning tool. By developing a data set and visualization of those who died at the hands of police in "The Counted," the *Guardian* created an open data set that was later used to show how governmental data was not tracking this problem properly.

Debate is important to our society. Open data is clearly political, and so too are data visualizations. Not only do they help to communicate insights found in data, but the very act of making them also belies an agenda. Whether it is to illustrate and persuade, as in Justice Sotomayor's dissent, or to argue for or against climate change, data visualizations set apart a topic or an idea as something of significance or importance. Data visualizations work so well

at communicating policy because they embody a sense of legitimacy, which can sway the public. This is of course a double-edged sword as ultimately the messages data visualizations communicate come from those who design them. Being transparent, involving the public, and collaborating in the creation of data visualizations helps, but as a society we need to become more data literate so that we can critique the messages data visualizations provide.