BIG DATA FOR CITIES IS NOT NEW

Big data for cities is not new. Some might argue that data collection is as old as civilization itself. In its earliest forms, data gathered about property rights, land production, and assets was used primarily to control and extract value by those in power. James C. Scott's book, *Seeing Like a State*, reminds us that units of measurement were developed to exploit resources. For example, landowners found ways to reduce or increase the size of a bushel of wheat—perhaps by specifying whether the grain was to be heaped or leveled off, or poured from shoulder or waist height (the "longer" the pour the tighter it packed)—and they did so in their own favor as they set the standard.[1] Throughout history we have seen data used to oppress and marginalize populations, sometimes unintentionally but other times on purpose. And yet historical accounts have also shown data analytics used to improve society when applied to everything from social services to public health and eradicating disease. Ultimately, the intent of the person analyzing the data affects how it is used. Applying data to city development has presented great benefits but also posed real dangers. For that reason (at least) data analytics for city planning sparks controversy.

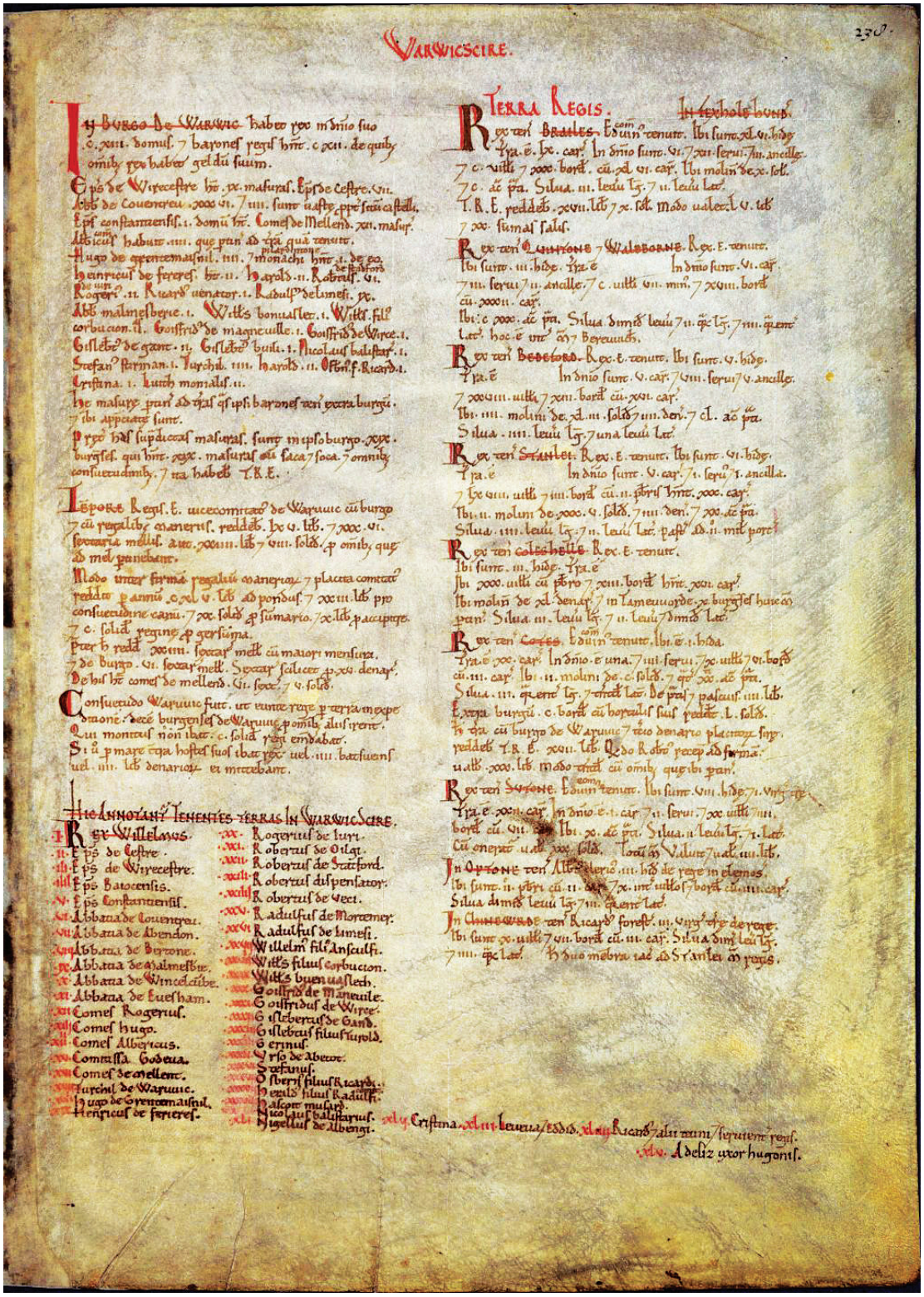
Data is essential for governance, of course, but in the hands of different people it will produce wildly different outcomes. This chapter provides a short history of data exploits in the city by primarily focusing on the use of data during and after industrialization, the period when society became increasingly fascinated with how scientific methods could be used to understand social processes. During industrialization we saw increases in

the use of the modern census and vital statistics (population counts, births, and deaths)—all of which, being essential to governance, became a point of political debate. But the complexity of cities was often oversimplified in data models developed in postwar America, and that often led to harmful decisions. Learning from the historical successes and failures of how cities have used urban data will equip us to apply data to changing policy more ethically and responsibly. This abbreviated history of data use in cities will not only help frame how to use data for action, but also show how data can be used to wield social and political control—with sometimes devastating, other times salutary, but nonetheless lasting effects.

Data: Essential for Governance

Governing bodies of the earliest civilizations gathered data about populations in order to provide services, build infrastructure, collect taxes, and enforce policy. Historically this data was only available to the ruling class or nobility to help control the populace.[2] Ancient Rome also carried out censuses (*census* is a Latin word) as early as the fifth century BCE, and the Romans are often cited as the first to undertake a social survey. Similarly, it has been said that the earliest dynasties (and the Chinese Empire itself) would not have survived without the large bureaucracy and data collection that enabled it to control its vast territories.[3] Not surprisingly, China is home to the oldest preserved census, which in 2 CE registered 57,671,400 individuals.[4] During the Middle Ages in England, William the Conqueror (King William I) ordered all property owned by his subjects to be recorded in the *Domesday Book* (1086)—not only for the purpose of raising taxes for his army but also to determine England's wealth (figure 1.1). For the most part, data continued to appear as text in published books until William Playfair's *Commercial and Political Atlas and Statistical Breviary* (1786) introduced the "universal language" of charts and pie graphs (see figure 4.1 in chapter 4).[5] In contrast, the Incas relied on a different visual method of communication, recording census data, history, and property records in *khipu*, elaborately tied strings made of camelid hair (figure 1.2).[6]

Overall, census data today is more openly available to the public than ever before. It is used routinely by governments around the world to make decisions about everything from taxation to representation in government, to the provision of social services.[7] Still, while taking a census seems



WARWICSCIRE. 238.

IN BURGO DE WARWIC habet rex in dñio suo .c. xxii. domos. 7 baronef regis hñt .c xii. de quib omib rex habet geld dū suum.

Eps de Wirecestre hē .x. masuras. Epf de Cestre. vii. Abb de Couentreu .xxx vi. 7 iii. sunt assise ppē suā castellā. Epf constantiensis. i. domū hē. Comes de Mellend. xii. masuras. Abbicia habuit .iii. que pau' ad tēra quā tenuit. Hugo de grentemaisnil. iii. 7 monachi hñt .i. de eo. Henricus de ferers. hē .ii. Harold .ii. Robertus. vi. Rogeri .ii. Ricard uenator. i. Radulf de limesi. ix. Abb malmesberie. i. Wills bonuslez. i. Will. filius corbucon. ii. Goisfrid de magneville. i. Goisfrid de Wirc. i. Gislebt de gant. ii. Gislebt buili. i. Nicolaus balistar. i. Stefan sarman. i. Turchil. iii. Harold. ii. Osm f. Ricardi. Cristina. i. Kirch monialis. ii.

Hę masurę pau' ad tēras ipsi baronef regis exeūt burgū. 7 ibi appendāt sunt.

Preter hās supdictas masuras. sunt in ipso burgo .xxx. burgenses. qui hñt .xxx. masuras sū saca 7 soca 7 omnib consuetudinib. 7 ita habebāt T.R.E.

Tempore Regis. E. uicecomitat de Warwic cū burgo 7 cū regalib manerijs. reddeb hē .v. lib. 7 xxx. vi. sextarios mellis. 7 ad .xxxiii. lib 7 vii. solid. p omib. que ad mel panebant.

Modo inter firmā regalū maneriorū 7 placita comitat reddit p annū .c xl v. lib ad pondus. 7 xx iii. lib pro consuetudine carū. 7 xx. solid p sumario. 7 x. lib p accupre. 7 c. solid regine p gersuma.

Preter h redd .xxiii. sextar mell cū maiori mensura. 7 de burgo .vi. sextar mell. Sextar scilicet p xx. denar. De his hē comes de mellend .vi. sext 7 v. solid.

Consuetudo Warwic fuit. ut tunc rege p terra mexpe dnone. decē burgenses de Warwic p omib alus irent. Qui nollet non ibat. c. solid regi emendabat. Si ū p mare exta hostel suos ibat rex uel .iii. barcuens uel .iii. lib denariof ei mittebant.

HIC ANNOTANT TENENTES TERRAS IN WARWICSCIRE.

1 Rex Willelmus.
2 Eps de Cestre.
3 Eps de Wirecestre.
4 Eps Baiocensis.
5 Eps constantiensis.
6 Abbatia de Couentrei.
7 Abbatia de Abendon.
8 Abbatia de Bercone.
9 Abbatia de Malmesbir.
10 Abbatia de Wincelcūbe.
11 Abbatia de Euesham.
12 Comes Rogerius.
13 Comes Hugo.
14 Comes Albericus.
15 Comitissa Godeua.
16 Comes de Mellent.
17 Turchil de Warwic.
18 Hugo de Grentemaisnil.
19 Henricus de Ferers.
20 Rogerius de Iuri.
21 Robertus de Oilgt.
22 Robertus de Stafford.
23 Robertus dispensator.
24 Robertus de Vecti.
25 Radulfus de Moreney.
26 Radulfus de Limesi.
27 Willelm filius Ansculfi.
28 Wills filius Corbucon.
29 Wills buenuaslech.
30 Goisfrid de Maneuile.
31 Goisfridus de Wirec.
32 Gislebertus de Gand.
33 Gislebus filius Turold.
34 Grimus.
35 Vrso de Abeton.
36 Stefanus.
37 Osberni filius Ricardi.
38 Herulf filius Radulfi.
39 Walter[illegible]
40 Nicolaus balistarius.
41 Nigellus de Albengi.
42 Cristina.
43 Leuuin/Edid.
44 Ricard alii xiii/seruiens regis.
45 Adeliz uxor hugonis.

TERRA REGIS. In texhole hund.

Rex ten Bratles. Edwin com tenuit. Ibi sunt .xl. v. hid. Tra. ē. hx. car. In dñio sunt .vi. 7 xii. seruī. 7 iii. ancille. 7 c. uilli 7 xxx. bord cū .xl. vi. car. Ibi molin de .x. sol. 7 c. ac pra. Silua. iii. leuu lg. 7 ii. leuu lat. T.R.E. reddeb .xxvii. lib 7 x. sol. modo ualet .l. v. lib 7 xx. sumar salis.

Rex ten Quinrone 7 Walborne. Rex. E. tenuit. Ibi sunt .iii. hide. Tra. ē In dñio sunt .vi. car. 7 iii. seruī 7 ii. ancille. 7 c. uilli 7 vi. mīr. 7 xxvii. bord cū .xxxii. car. Ibi .c xxx. ac pra. Silua dimid leuu 7 ii. qr lg. 7 iii. qr em lat. hoc ē int ām 7 Bereuuod.

Rex ten Bedeford. Rex. E. tenuit. Ibi sunt .v. hide. Tra. ē In dñio sunt .v. car. 7 vii. seruī 7 v. ancille. 7 xxvii. uilli 7 xvi. bord cū .xxi. car. Ibi .iii. molini de .xl. iii. solid 7 iii. den. 7 cl. ac pra. Silua .iii. leuu lg. 7 una leuu lat.

Rex ten Stanlei. Rex. E. tenuit. Ibi sunt .vi. hide. Tra. ē In dñio sunt .v. car. 7 i. seruus 7 i. ancilla. 7 lx. vii. uilli 7 iii. bord cū .ii. pbris hñt .xxx. car. Ibi .ii. molin de .xxx. v. solid 7 iii. den. 7 xx. ac pra. Silua .iii. leuu lg. 7 ii. leuu lat. past ad .ii. mil pore.

Rex ten Coleshelle. Rex. E. tenuit. Ibi sunt .iii. hide. Tra. ē Ibi .xxx. uilli cū pbro 7 xvii. bord hñt .xxi. car. Ibi molin de .xl. denar 7 in Tameuuorde .x. burgenses hanc ām pan. Silua .iii. leuu lg. 7 ii. leuu 7 dimid lat.

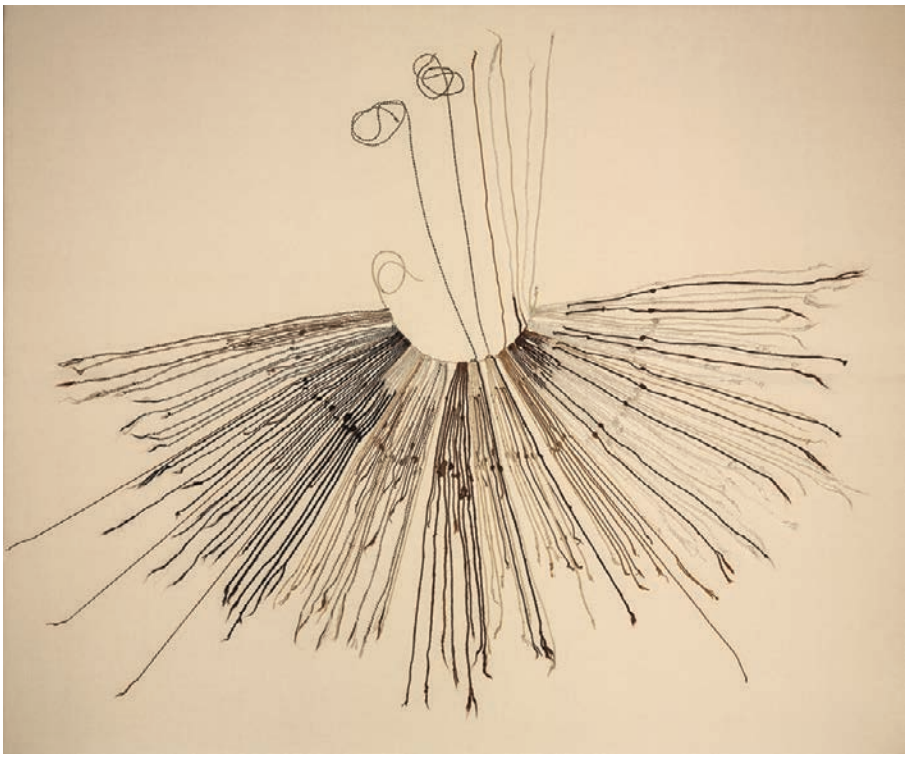
Rex ten Cotes. Edwin com tenuit. Ibi ē .i. hida. Tra. ē .xx. car. In dñio ē una. 7 iii. seruī. 7 xx. uilli 7 vi. bord cū .iii. car. Ibi .ii. molini de .c. solid. 7 qr .xx. ac pra. Silua .iii. qr em lg. 7 dim lat. De pastu pasaus .iii. lib. Extra burgū .c. bord cū boscaul suis pedar .l. solid. H tra cū burgo de Warwic 7 duo denarii placent sup. reddeb T.R.E. xvii. lib. Q.do R. obt recep ad firmā. ualb .xxv. lib. modo dim cū omib que ibi pan.

Rex ten Sutone. Edwin com tenuit. Ibi sunt .viii. hide. 7 i. virg tre. Tra. ē .xx. car. In dñio ē .i. car 7 ii. seruī. 7 xx. uilli 7 iii. bord cū .vii. car. Ibi .x. ac pra. Silua .ii. leuu lg. 7 i. leuu lat. Cū emendō ualb .xx. solid. Tot ām Valuit 7 ualet .iii. lib.

In Optone ten Alric de rege in elemos. Ibi sunt .ii. pbri cū .ii. car 7 .x. mr uilli 7 bord cū .iii. car. Silua dimid leuu lg. 7 iii. qr em lat.

In Cumendune ten Ricard forest. iii. virg tre de rege. Ibi sunt .x. uilli 7 vii. bord cū .iii. car. Silua dim leu lg. 7 iii. qr lat. H duo mebra iac ad Stanlei ām regis.

1.1 A page from the *Domesday Book* (1086) listing land holdings in Warwickshire. Source: courtesy of J. J. N. Palmer, "Warwickshire, page 1," Open Domesday, n.d., https://opendomesday.org/book/warwickshire/01/.



1.2 Incan fiber recording device (*khipu*) found in Peru, c. 1400–1532. Cords, knotted and twisted; cotton and wool. Overall: 85 × 108 cm (33 7/16 × 42 1/2 in.). Source: Cleveland Museum of Art, https://clevelandart.org/art/1940.469.

fundamental to governance, many countries do not conduct such surveys. Some researchers, for example, suggest that fewer than half the births in Africa in 2014 were recorded.[8] This makes planning for infrastructural improvements extremely difficult in many regions, for reasons that range from a particular government's insufficient resources to its lack of interest.[9] This is why one of the most recent United Nations Sustainable Development Goals (SDGs) includes the development of data—they want to encourage governments to collect it.[10]

The Census: Contesting Democracy through Data

Census data is a political tool. It is used to plan infrastructure, understand citizens' needs, track economic development, and more generally measure the economic and social health of communities. In *The American Census: A Social History*, Margo Anderson reminds us that the census reflects American values and the concerns of the time.[11] Every ten years, questions are added to track real and emergent trends, such as employment status, language spoken at home, or access to the internet.

Census data allows us to grasp some of the biggest unknowns in a society in order to address them through public policy. As the United States industrialized in the nineteenth century, the office of the census was professionalized. It used new statistical techniques that allowed for data sampling; and it relied on tabulating and punch card machines—invented by a census bureau worker who later founded the company that would eventually become IBM—to count responses. This was all needed to track the unprecedented growth of a country expanding in all directions. During this time, the Census Bureau also added questions about cities to reflect societal changes that were occurring as their populations doubled.[12] Census questions, whether additions or removals, tell us what governments and the public are interested in measuring, and as such reflect the values of a society.

Redistricting: The Politics of Inclusion and Exclusion

America's congressional redistricting processes, which involve redefining geopolitical boundaries, are one example of how a seemingly benign data set—the census—can produce varying outcomes, depending on the values of those who is applying the census's data. The US census was established to apportion congressional representation and determine taxes. From the very first census in 1790, the question of who should be counted caused political debate. For example, enslaved people were not recorded as full individuals but rather at three-fifths of their total. It was not until the ratification of the Fourteenth Amendment in 1868 that African Americans were counted *fully* as part of congressional redistricting.

The politics of inclusion and exclusion in the census continues today. The Trump administration proposed to include a citizenship question on the 2020 census that opponents claim will severely undercount the population,

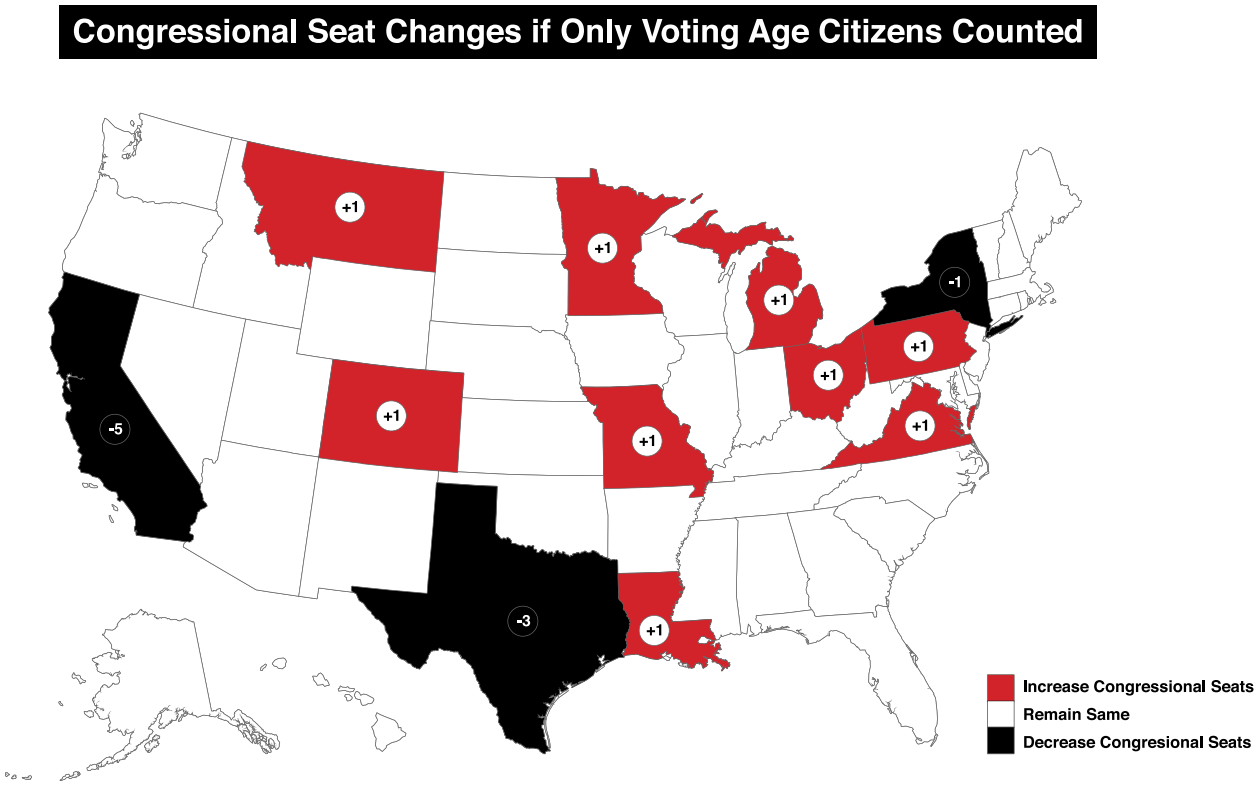
particularly in cities or areas with high numbers of immigrants who might be reluctant to be recorded.[13] According to a *New York Times* article, "if 15 percent of noncitizens went uncounted, that would be enough to cost California and New York one congressional seat each, to the benefit of Colorado and Montana."[14] The picture would be even more dramatic if only voting-age citizens were counted: California would lose five seats (figure 1.3). Government funding for services would also be affected, since federal funds are often allocated based on population figures. As of November 2018, more than two dozen states and cities had filed lawsuits against the Trump administration to get the question removed, citing that simply asking whether you are a citizen could have serious political consequences.

As of June 2019, the citizenship question was still under debate as documents emerged from the hard drive of the late Thomas B. Hofeller, a Republican operative whom the *New York Times* called the "Michelangelo of gerrymandering."[15] The documents show that Hofeller had performed a study in 2015 detailing how adding certain census questions would be to the Republicans' advantage.[16] One memo quotes Hofeller saying that adding the question regarding citizenship could be "advantageous to Republicans and non-Hispanic Whites," implying that it would reduce the number of residents who were tallied in the census from typically Democratic-voting districts with large immigrant populations.[17] This is perhaps the most current illustration of how decisions regarding data can be used to the advantage of one political ideology and party over another.

Hofeller dedicated most of his career to using data to change the political map from blue to red. In 1970, with the release of the first digital census, algorithms were used to redistrict, and Hofeller was one of the first people to apply computation models. Considered a data geek by his colleagues, he cynically supported black redistricting efforts under Section 2 of the Voting Rights Act. His goal was not to advance equal representation, something that generations of African Americans had been fighting for, but rather to clear African Americans—the most reliable Democratic voters—out of Democratic districts. Perhaps the most striking example of his intent was discovered in files found on his computer after his death that showed he used extensive profiling data to redraw congressional district lines directly through North Carolina Agricultural and Technical State University in Greensboro, North Carolina. The result divided this historically black college's campus in half, incorporating each side into majority Republican districts and

thereby diluting the political voice of the A&T students. These congressional districts have since been found unconstitutional, using Hofeller's own notes as evidence.[18] Hofeller has been implicated in numerous gerrymandering cases. According to his obituary in the *New York Times*, Hofeller's creative redistricting exercises were responsible for shifting the US Congress to Republican control in 2010.[19]

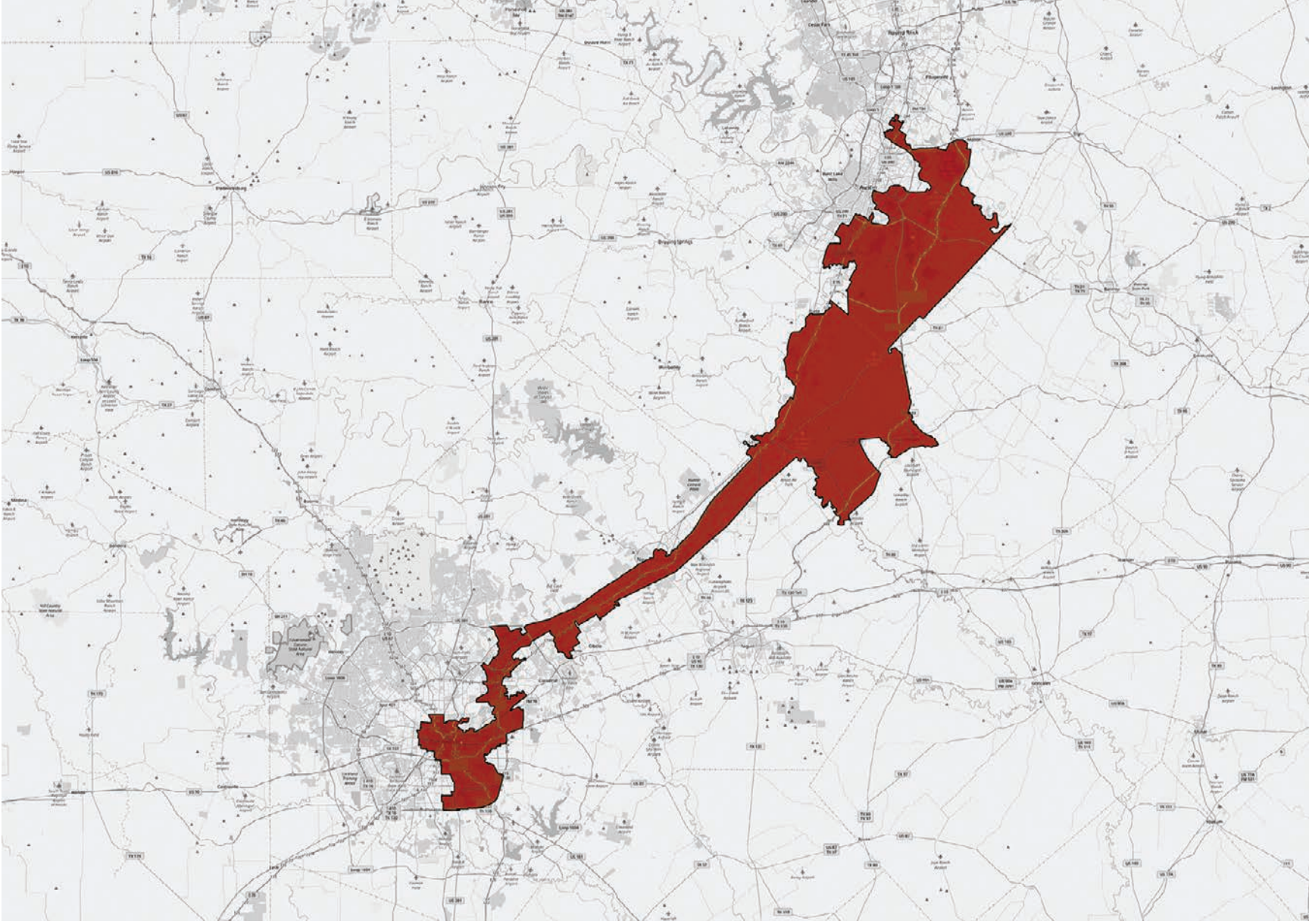
Hofeller's work benefited from the lack of standardization in the redistricting process state to state, making it easy to manipulate the data toward the will of each state's redistricting committee. Officially, congressional



1.3 The map shows congressional seats that might be lost or won if an "only citizens of voting age" qualification were to be included in determining how congressional districts are drawn. Right now, we use the total population of each state, regardless of citizenship or eligibility to vote, to determine congressional seats. Source: Created by the Civic Data Design Lab at MIT, using data from the 2016 American Community Survey.

districts must reflect equal populations, minority representation, contiguity, and compactness, and each district must try to maintain existing political or geographic boundaries. But the decisions about how to interpret these rules rest with committees in each state whose members ultimately approve the new districts. Some states seek to preserve communities of interest (whether social, cultural, ethnic, economic, religious, or political), whereas others attempt to maintain current electoral seats.[20] But this leaves room for bias, because each state can decide how it develops its redistricting commissions. Some appoint bipartisan commissions while others select members of the majority party. States where redistricting is performed by the party in power (in Texas, for example) often see partisan gerrymandering.[21] Other states are totally inclusive (New Hampshire is one), allowing their whole legislature to make redistricting decisions. Once legislatures have redrawn congressional districts, only a court case can change any contested boundaries—and that can be a long and costly process.[22] Ultimately, redistricting shows how pushing around data can have lasting political consequences.

Fear of immigrants, represented by the emergence of the citizenship question in the 2020 census, is well documented in US political history, and according to some represents a longstanding divide between urban and rural America. In his book *Delayed Democracy*, Charles W. Eagle describes how people in the 1920s perceived rural areas as "strongholds of the Klan, prohibitionism, fundamentalism, and nativism. . . . The cities, by contrast, represented tolerance, progress, and aspects of American life." This division between urban and rural was so strong, Eagle writes, that when the 1920 census data showed for the first time how urban areas had grown larger that rural areas, the redistricting process was brought to a standstill because many rural areas feared a loss of power. It was not to their advantage to redistrict because they would surely lose congressional seats to urban areas. The standstill in redistricting meant that between 1920 and 1930, congressional districts were widely unbalanced, effectively muting this new urban population's political voice. The imbalance was eventually stabilized by the Reapportionment Act of 1929, which mandated states to redistrict in a timely manner, but questions about how to redraw district lines have remained a point of contention between urban and rural communities. For instance, in a 2017 ruling, a federal court found the redrawing of three congressional districts in Texas, including the 35th (figure 1.4), to be unconstitutional.[23] These lines were drawn to help retain congressional seats held by Republicans, who are often associated with rural America.



1.4 In 2017 a panel of federal judges ruled that the 35th Congressional district in Texas violated the Constitution and Voting Acts rights because race had been a motivating factor in redistricting that occurred in 2011. This district has become synonymous with gerrymandering. Source: https://www.npr.org/sections/thetwo-way/2017/03/11/519839892/federal-court-rules-three-texas-congressional-districts-illegally-drawn.

Planning for the Unknown City: Data and the Industrializing City

The rural and urban divide started when industrialization caused people to flock to cities from the countryside. This dramatic change in the character of cities created vast contrasts between the life of the rich and poor: during this time we see the emergence of vast slums and tenement housing. In many

cities, these transformations put wealthy, white elites in direct contact with poverty, making them uncomfortable. At the same time, a new professional class (including medical doctors and clergymen) now had growing access to data that was once only available to government administrators.[24] Relying heavily on vital statistics available through the census, surveys, mapping, and qualitative images and interviews, some Victorian-era reformers used this new access to data in order to develop statistical techniques that would help the poor. Their work gave birth to the field of public health and the development of modern sanitation. Their outputs were often highly visual: graphs and maps helped advocate for their causes.

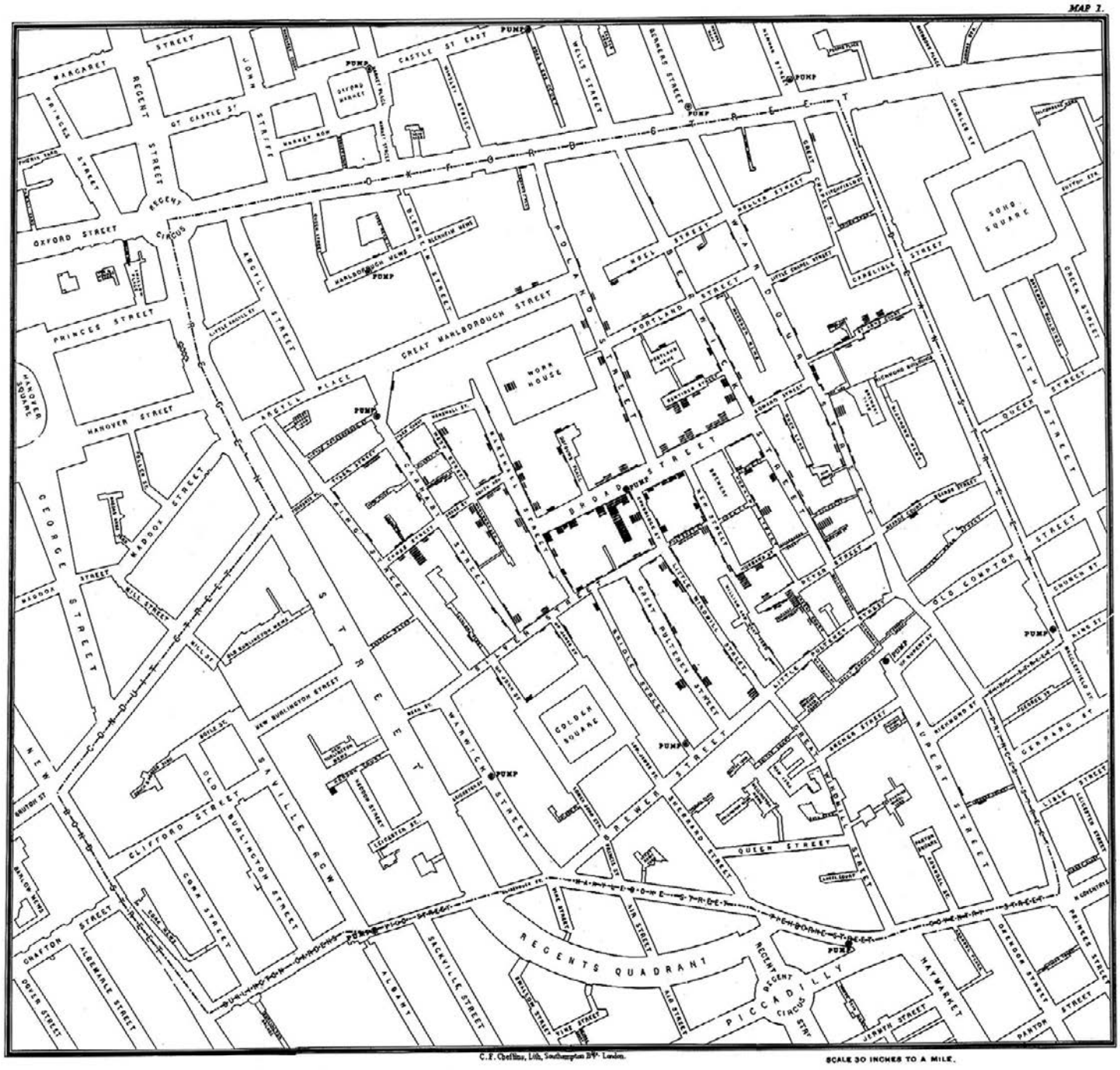
Data Analytics: Giving Birth to the Field of Public Health

The First Industrial Revolution (1760–1840) transformed the dependency on manual labor in the economies of societies to the deployment and reliance on steam-powered machines. Urban populations grew tremendously as people moved from the countryside seeking work in factories. London, perhaps the first mega-city, doubled its population between 1802 and 1850. With that growth came extreme poverty, a lack of proper housing and sanitation, and outbreaks of cholera (in 1832 and 1850s).[25] Such outbreaks, it turns out, propelled early nineteenth-century data collection efforts, often in direct response to public health issues, living conditions, or poverty.[26] In Britain, this interest resulted in the establishment of a number of statistical societies,[27] the first of which, the Manchester Historical Society, was founded in 1833, at approximately the same time that the United Kingdom's Ordnance Survey started to create cadastral (detailed map) surveys of cities.[28] Just a year later, in 1834, Charles Babbage co-founded the Statistical Society of London, which later became the Royal Statistical Society. Increased access to data helped encourage the development of these new organizations to address issues important to society.

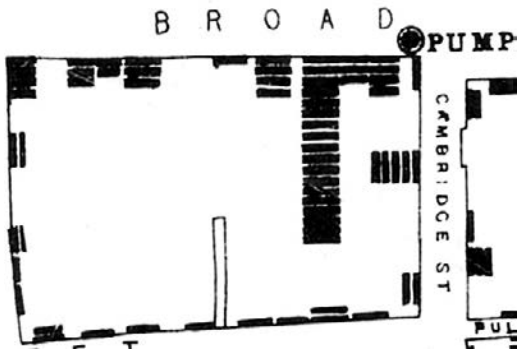
The fields of epidemiology and public health were born from this new, heightened interest in statistical analysis and mapping. Edwin Chadwick, John Snow, and William Farr, among other notable public health statisticians, were members of these societies. Their sharing of data and their analytical methods generated groundbreaking research and helped eradicate disease. Most notably, perhaps, they shared their data through visualizations that communicated their ideas clearly enough for the results to be effective at

1.5 Sanitary map of the Town of Leeds from the report of Sanitary Condition of the Labouring Population of Great Britain (1842). Source: "1842 Sanitary Report," https://www.parliament.uk/about/living-heritage/transformingsociety/livinglearning/coll-9-health1/health-02/1842-sanitary-report-leeds/.

changing policy. Chadwick developed the *Report on the Sanitary Condition of the Labouring Population* (1842), which graphically described the physical and social concerns that might be contributing to outbreaks of cholera (figure 1.5). The compelling findings and the ease with which the data was communicated through maps generated debate that is said to have helped lead to the passage of Britain's Public Health Act in 1848.[29]



1.6 John Snow's cholera map (1854) shows the high numbers of deaths from cholera around the Broad Street water pump. The map is often considered an example of data visualization because it included bar charts to pinpoint the number of deaths at a particular address. Source: John Snow, *On the Mode of Communication of Cholera* (London: John Churchill, 1855).



1.7 Zoom of the bar charts used on John Snow's cholera map (1854), shown in figure 1.6.

John Snow developed his now famous map (figures 1.6 and 1.7) of the Broad Street Pump (1855), which linked deaths from cholera to the contaminated pump water. This could only be done after William Farr, Compiler of Abstracts (or chief statistician) for Britain's General Register Office (GRO), openly shared with Snow statistical data on deaths from cholera. Snow's maps showed the geographic concentration of cholera deaths to those households who used the Broad Street pump as a water source.[30] Before Snow's mapping experiment, scientists did not widely accept that cholera could be carried in water, but the maps provided overwhelming visual evidence of the need to revise this assumption. Snow's visualizations have had a lasting influence on the fields of public health and sociology because of their ability to communicate information essential to policy needs and policy-making. Farr, who was the equivalent of a present day chief technology officer, went on to develop the statistical evidence for Snow's maps, showing that his work was neither an anomaly nor visual artifice.

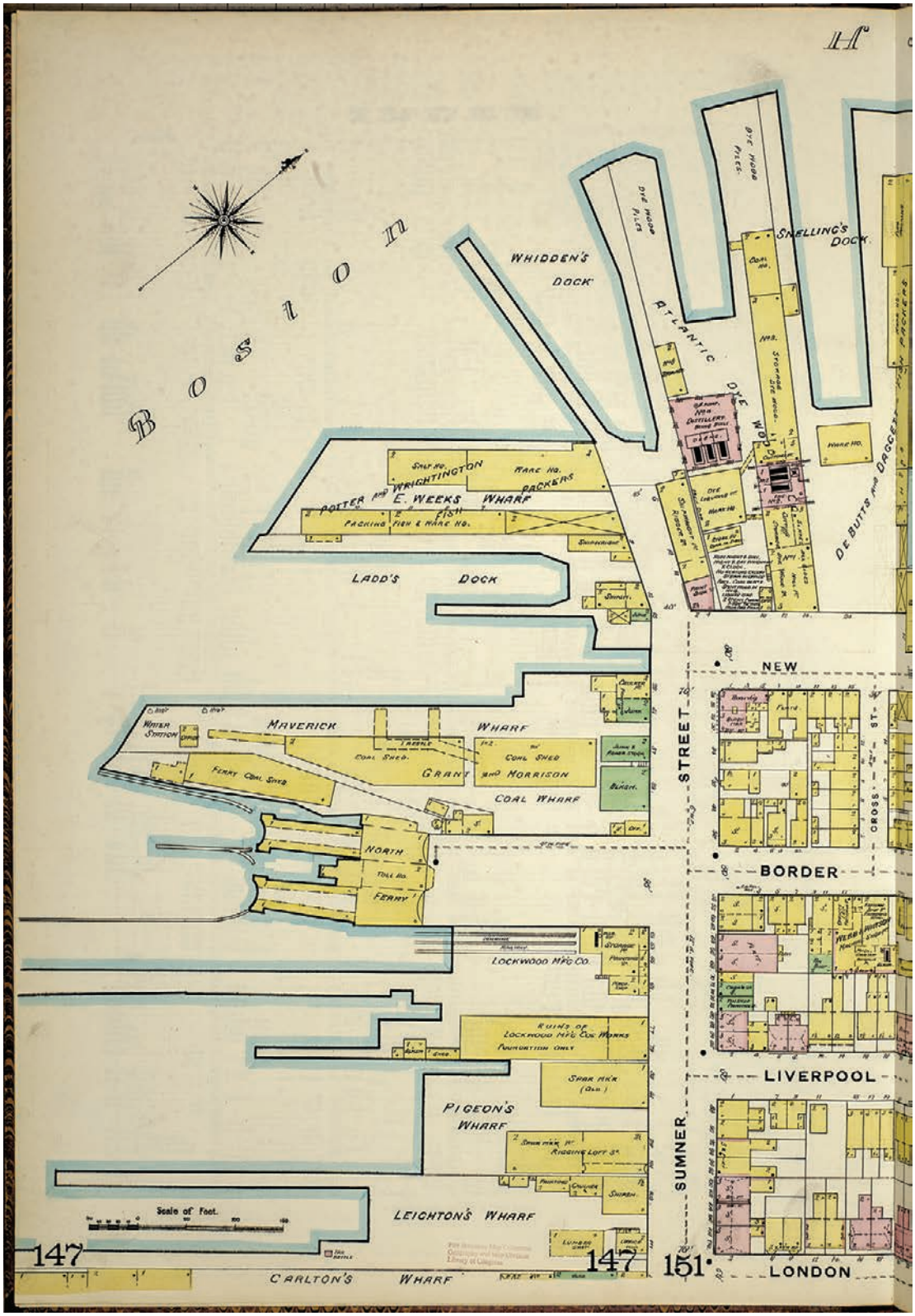
Farr's tenure (1851–1871) overseeing Britain's censuses and his role as Britain's chief statistician might be considered the first open data gesture: it marks an important, early moment in the growth of Britain's open data movement. Under his direction, the types of data available in England multiplied and began to include such vital statistics as place of birth and registration of deaths (1841).[31] Farr performed extensive studies using death certificates, looking at the rates of suicides among different classes and education levels, occupational mortality, and causes of death. As he became interested in the spread of disease he studied influenza (1847) and cholera (1854) in particular.[32]

Cadastral Mapping: Controlling the Industrialized City

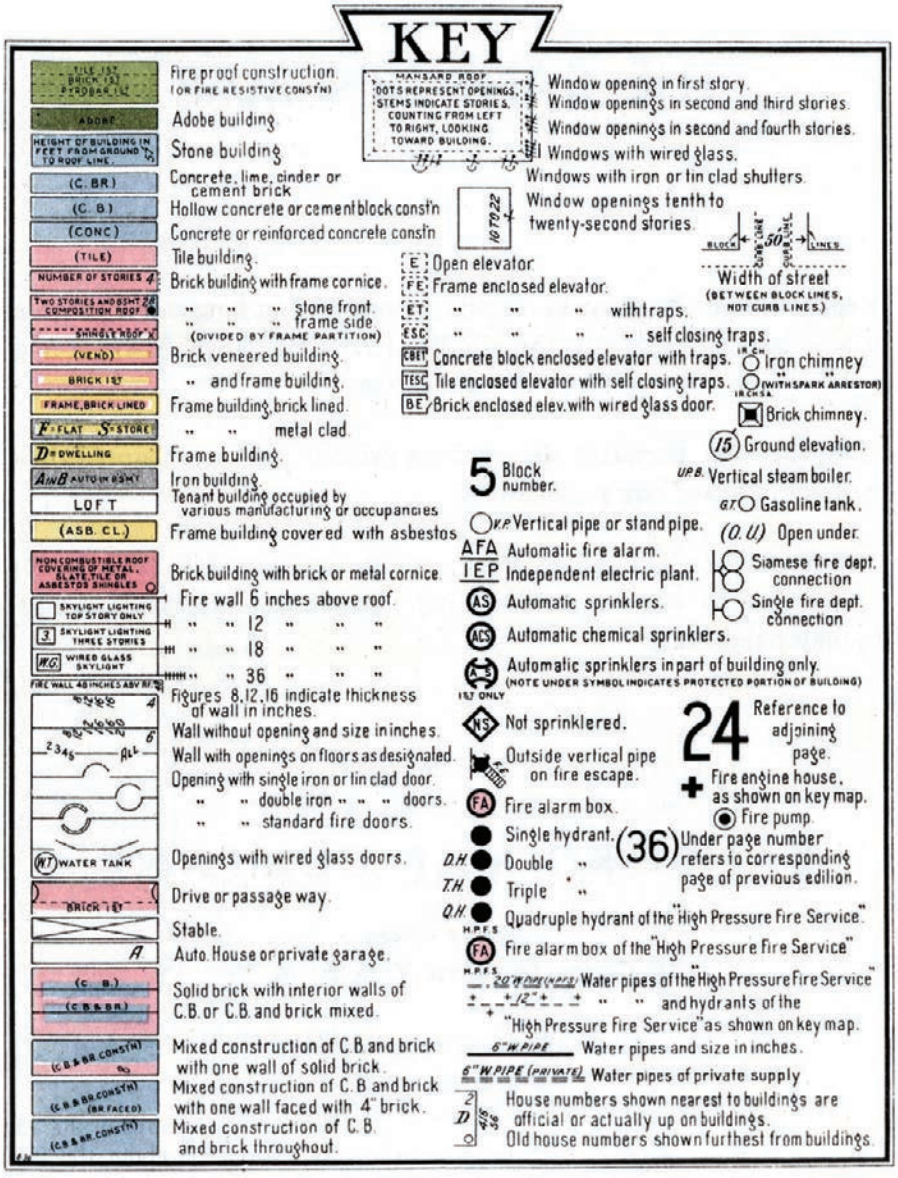
Data in the form of maps became a communication tool for Victorian-era social statisticians to document land assets, poverty, and "subversive elements" of the city. By highlighting details such as the locations of brothels or of immigrants with distasteful social norms, these maps became a tool for influencing policies to control unsavory elements of the city. Although many of these maps were conceptualized to address the plight of the poor, once put to use they often further marginalized them, simply by marking them as "other" on the map. These early maps illustrate the duality that exists within the application of data for policy change. Data can be used as evidence that can help or harm.

Victorian-era fascination with maps can be seen as early as 1824, when Britain's Ordnance Survey began developing detailed maps of cities and towns, showing for the first time building footprints, streets, roads, parks, and civic institutions such as churches, schools, and government buildings. These maps were largely used for taxation purposes. In 1862, the Ordnance Survey released one of the largest and most detailed in this series: London at a scale of 1:10,560 (where 6 inches on the map represents a real-world distance of 1 mile).[33] In the United States, private companies made similar maps; for instance, in 1867, the Sanborn National Insurance Diagram Company (later called Sanborn Fire Insurance) released its first map of Boston, which included incredible detailing of brick and wooden structures, contaminants, and types of businesses. The company went on to develop the same maps for most American cities. These early maps are still used today for environmental remediation because they include important data about where oil or other chemicals were used, thereby helping to construct an environmental history of cities (figures 1.8 and 1.9). Employing an army of cartographers for more than one hundred years, the Sanborn Company became the only source of data that traced the vast and rapid development of US cities.

Sanborn maps and their London cousins, the Goad Fire Insurance Plans (1886–1930), became an invaluable resource for land use information. Because they were focused on the built environment, however, the maps did not measure socioeconomic conditions of the city. Charles Booth, a London philanthropist, set out to change that. He believed a survey of poverty on the scale of these maps would be essential to study the condition of the poor.[34] This led Booth to conduct the most extensive survey on race and class London



1.8 This Sanborn map from 1888 shows a block in a manufacturing district in Boston, Massachusetts, including a coal shed, dye factory, and machine shops. The maps are an example of some of the first detailed data sets on cities. Source: Sanborn Map Company, "Image 3 of Sanborn Fire Insurance Map from Boston, Suffolk County, Massachusetts," 5 (1888), http://hdl.loc.gov/loc.gmd/g3764bm.g03693188805.



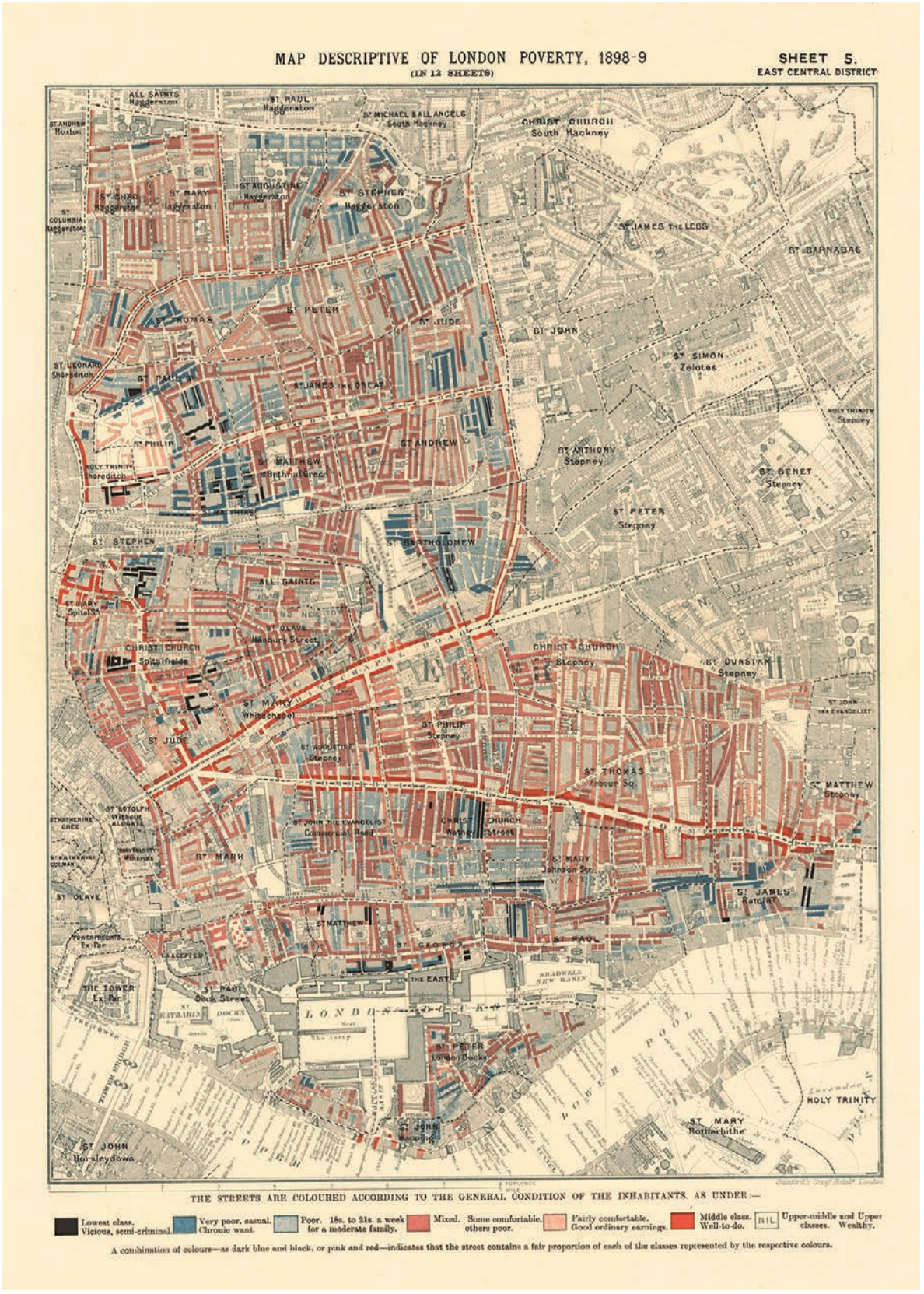
1.9 A typical key to a Sanborn map. The amount of detail is more than cities collect about most properties today. Source: FIMo—How to Interpret Sanborn Maps, 2019, http://www.historicalinfo.com/fimo-interpret-sanborn-maps/#Utilization.

had ever seen. From 1892 to 1903, Booth's team went to every property in London and asked a series of questions about income, quality of housing, and access to water, among other conditions of the urban environment.

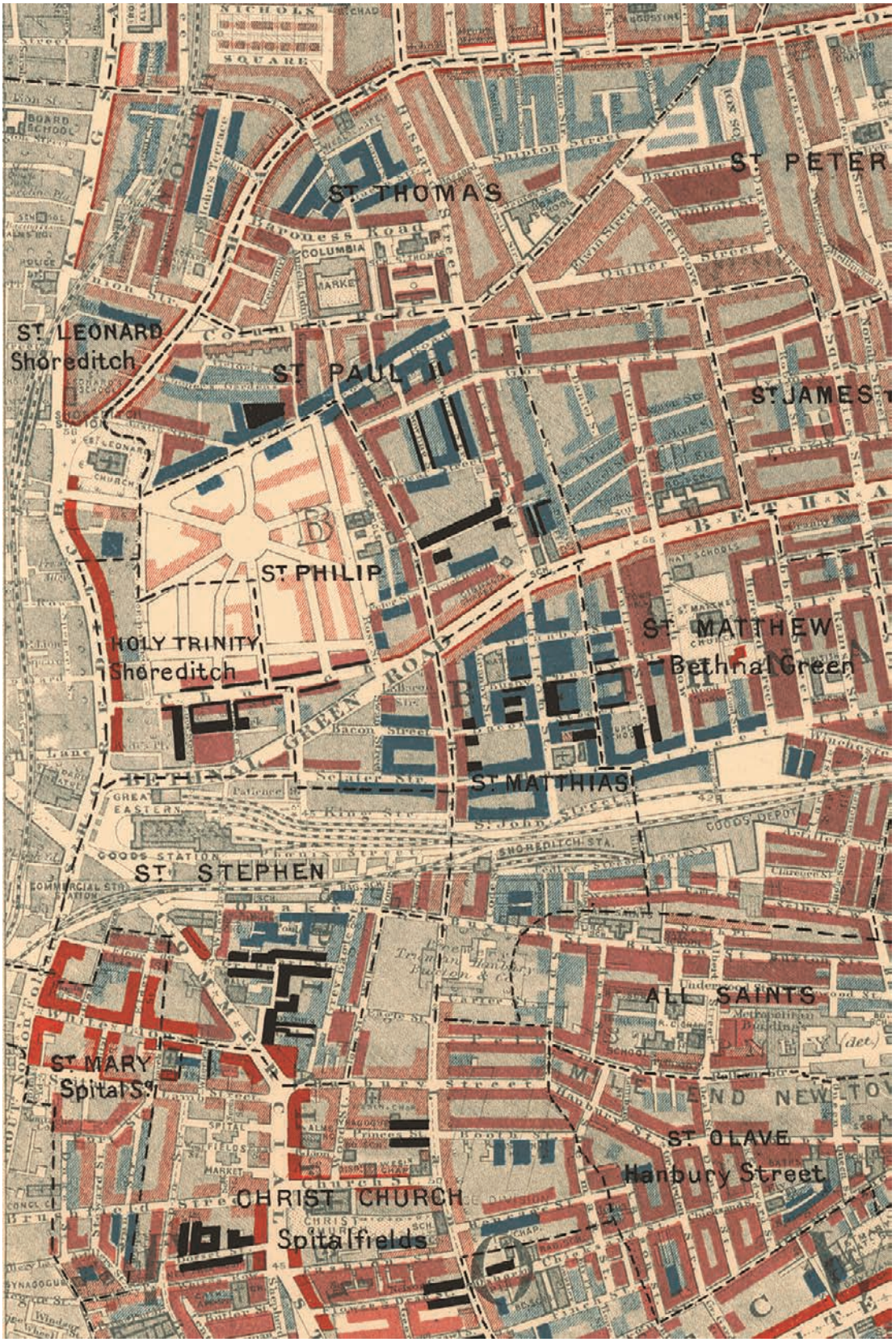
Booth's survey was translated into the now-famous *Descriptive Map of London Poverty* (figures 1.10a and 1.10b), which comprised the first ever sociodemographic maps of London.[35] Booth's map helped to properly estimate the poverty level, claiming that 30.7 percent of London's population lived in poverty.[36] Previous estimates were closer to 25 percent, so this represented a marked difference.[37] *The Maps* were considered exceptional insofar as they were the first to link social issues with the physical city; Booth's work helped philanthropists and governments to understand many social–spatial aspects of the lives of the urban poor in new ways (figure 1.10).[38] Although social surveyors of the time often employed tables, numbers, and statistics, Booth's maps were cited as proffering objective truth, adding credibility to his findings.[39]

At the same time, Booth's work became a tool to advance the systemized removal of whole populations from London's inner core; for example it inspired the development in 1901 of the Royal Housing Commission, where the maps were used to rationalize the complete demolition of slums and the rebuilding of social housing in their place. Nichol housing (a slum in Jewish East London), for example, was replaced by Boundary Estates, the first state social housing in Britain.[40] A housing upgrade may indeed have been needed in East London, but those who lived in the Nichol slum could not afford the rents for the new estates, which were set higher than what they had paid for Nichols.[41] As a result, an entire Jewish community was pushed out of their neighborhood. Looking back at the records, it is difficult to determine whether this was a purposeful targeting of a particular group. According to some accounts, Booth was "confused" by the Jewish culture and in some of his writings appears to be anti-Semitic.[42] This possible prejudice may have led to his focus on the Nichol area, although it appeared to have been slated for redevelopment before Booth's maps were available. Therefore, the maps may have simply created rational evidence to garner public support for policies the city government already wanted to enact.

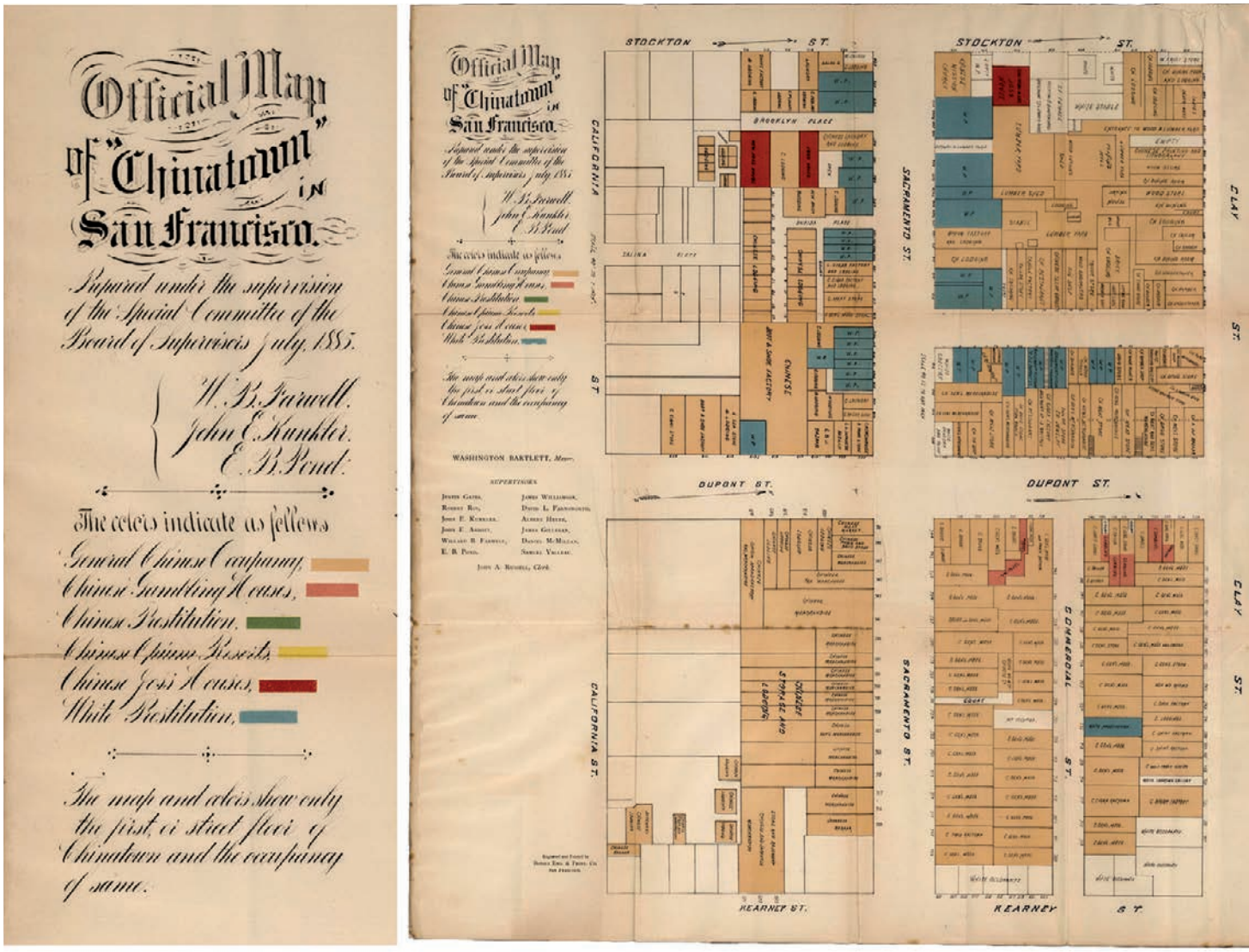
Philanthropists in America and Europe imitated Booth's work, and although many of the maps generated evidence for projects that helped the urban poor, others often had racist undertones and were used as evidence to enact discriminatory policies.[43] For example, in1885 San Francisco's Board of



1.10a and 1.10b Charles Booth's Social Demographic Maps of London. Figure 1.10a shows the original map sheet 5, focusing on East London, which was home to many Jewish communities at the turn of the twentieth century. Figure 1.10b is an enlarged view of the same map focused on the former location of St. Nichol slum, which was cleared to create Boundary Estate, one of London's first social housing experiments. Source: London School of Economics and Political Science, "Charles Booth's London," accessed August 2, 2019, https://booth.lse.ac.uk/map/13/-0.1565/51.5087/100/0.



Supervisors released a map of Chinatown's gambling, drug, and prostitution businesses (figure 1.11). The report that accompanied the map, which was written in the course of developing policies that would remove Chinese residents from the area, states: "We have permitted the Chinese to become our masters, instead of asserting and maintaining the mastery ourselves."[44] As told in Nayan Shah's book *Contagious Divides*, "nineteenth-century San Francisco health officers and politicians conceived of Chinatown as the preeminent site



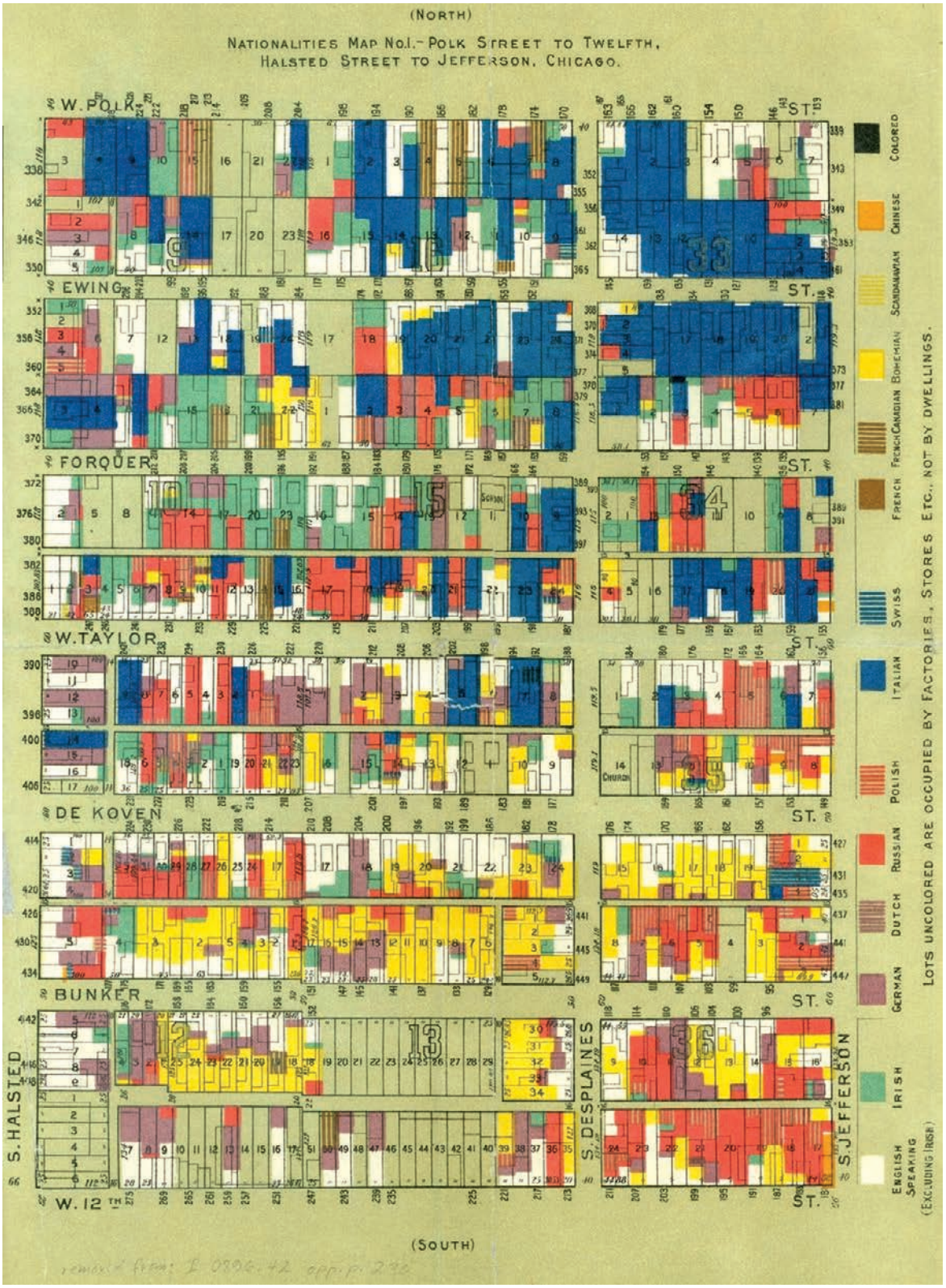
of urban sickness, vice, crime, poverty, and depravity."[45] Shah interprets the maps (figure 1.11) as surveillance, and their continued existence ensured this type of surveillance in the future. Maps became important tools for enacting city policies and were oftentimes used to segregate populations seen as "other." Where some employed cadastral maps to marginalize populations, many used them to expose and improve the quality of life of the urban poor. These types of cadastral maps were also used by the "settlement movement"

1.11 Maps from 1885 highlight businesses in Chinatown, specifically those considered unseemly, such as gambling, drugs, and prostitution. Source: "Mapping Vice in San Francisco," *Mapping the Nation* (blog)," accessed June 29, 2019, http://www.mappingthenation.com/blog/mapping-vice-in-san-francisco/.

to help the urban poor. Chicago's Hull House is perhaps the most famous of these settlement houses, which provided services such as childcare, education, and healthcare. In 1895, Jane Addams, who ran Hull House, and her partner Florence Kelley used Booth's London project as a model for the *Hull House Maps and Papers*. The maps helped to show that rents were not at all proportional to wages—helping to expose how landlords preyed on the urban poor (figure 1.12).[46]

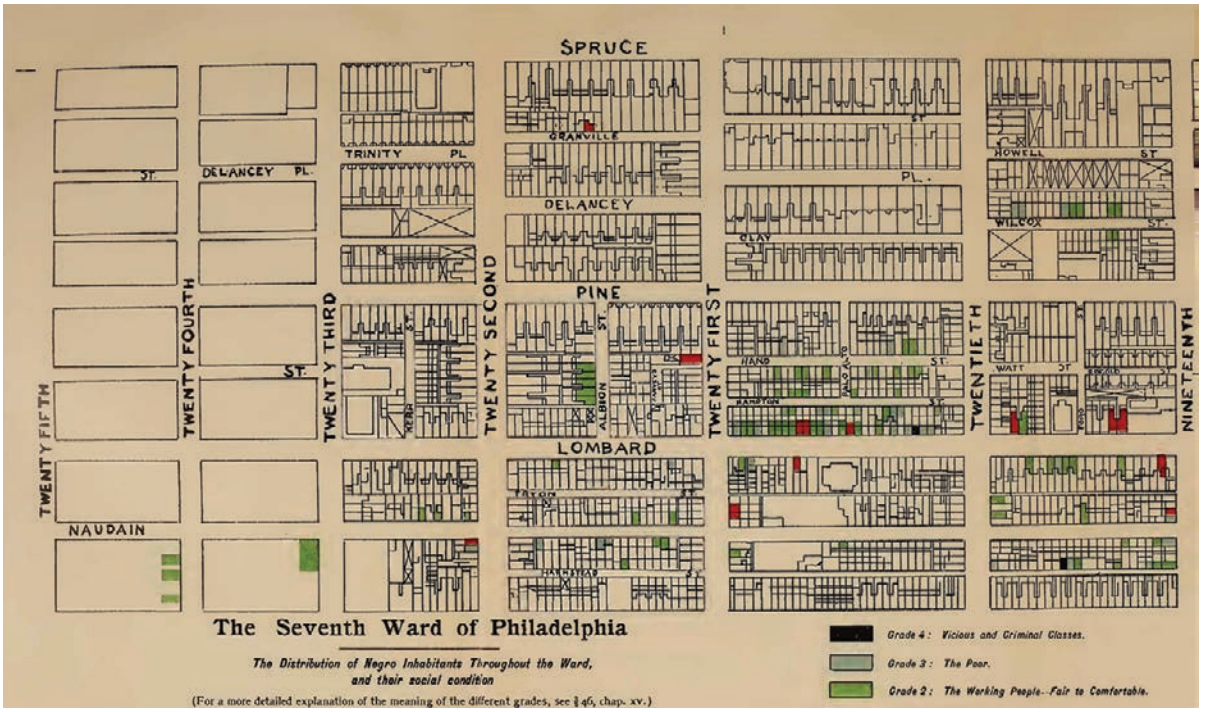
Addams and Kelley's work in turn influenced W. E. B. Du Bois's well-known study of the African American population in Philadelphia, released as *The Philadelphia Negro* (figure 1.13).[47] Du Bois extensively surveyed more than 5,000 residents regarding wages, employment, education, and housing stock, among other things. He noted that many African Americans lived in conditions that made their rise out of poverty difficult (figure 1.14).[48] Du Bois used the strategy in his other works, including his studies for the Conferences of the Atlanta Negro (figure 1.15).[49] In both Addams's and Du Bois's work, data visualizations in the form of maps helped create evidence for the existence of structural issues surrounding poverty, allowing viewers to reflect on those issues. How could an individual be expected to rise out of poverty when their wages were lower than even the most basic living expenses? Early reformers helped to raise awareness of this problem among broader groups of people.

It should be noted that at the same time that Jane Addams and W. E. B. Du Bois were working in Chicago, academics at the University of Chicago were also compiling qualitative data on the urban poor, eventually forming what became known as the "Chicago School" of sociology (1915–1935). Led by Robert E. Park and Ernest Burgess, the Chicago School saw the city as its urban laboratory and used the same methods of participatory observation that anthropologists were beginning to use with remote, indigenous populations.[50] The Chicago School's methods—which included data analytics, mapping, and qualitative research (figures 1.16 and 1.17)—were not meant to solve the problems of the city, but to try to define and describe the modern city, empirically, creating a "science of society." Their extensive mapping exercises led Burgess to create his famous "concentric zone" diagram (figure 1.16), which he used to explain racial segregation as a natural progression of city development, and therefore something that was unchangeable. These theories stood in stark contrast to how Jane Addams and W. E. B. DuBois used the same mapping techniques as tools to change economic and labor policy. The maps of the Chicago School are controversial because scholars believe

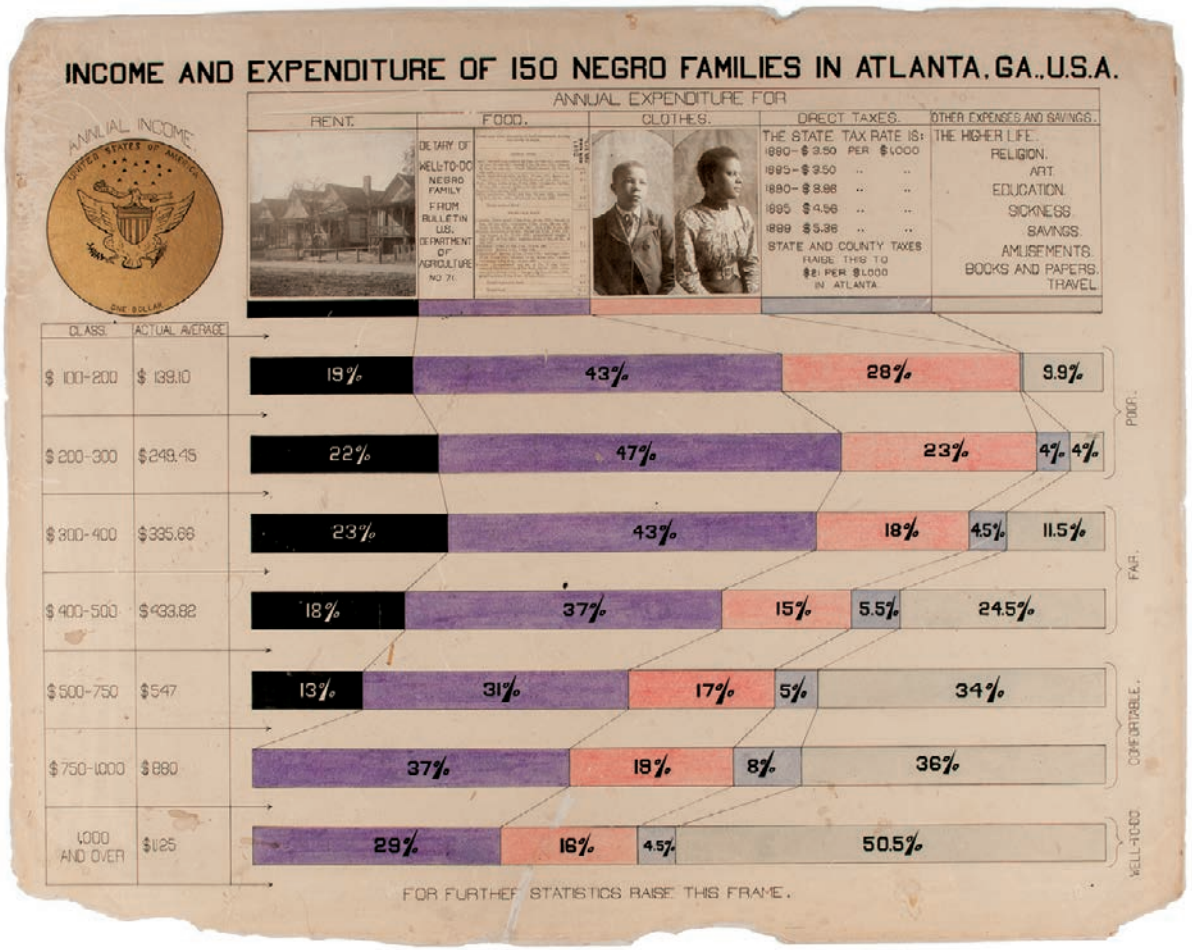


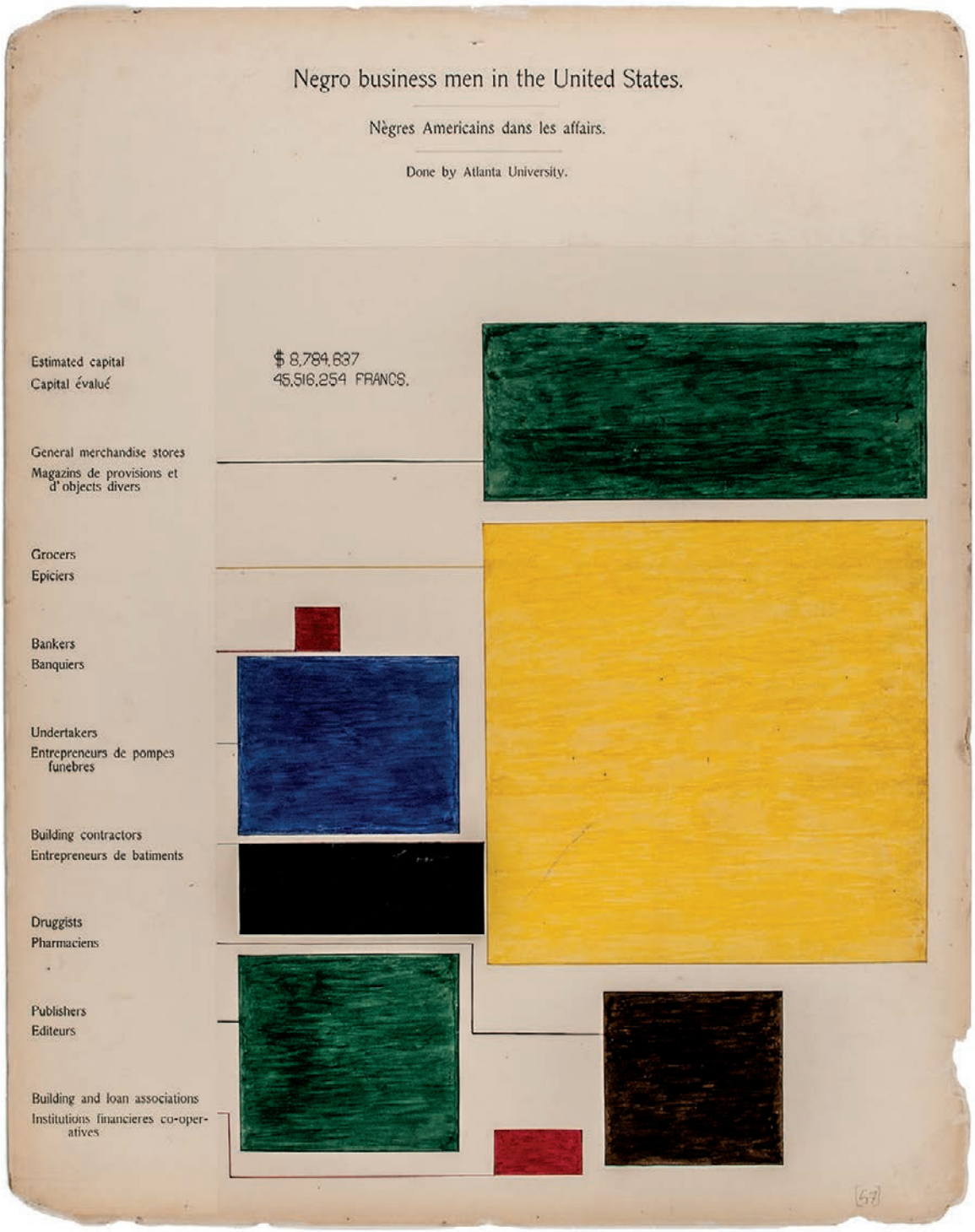
1.12 Map from Hull House Maps and Papers (1895) shows the different nationalities in an immigrant neighborhood around Hull House. Source: "Maps of Gilded Age San Francisco, Chicago, and New York," *Mapping the Nation* (blog), accessed June 29, 2019, http://www.mappingthenation.com/blog/maps-of-gilded-age-san-francisco-chicago-and-new-york/.

1.13 W. E. B. Du Bois included this map in The Philadelphia Negro (1899), an extensive survey of African American in the city's seventh ward of Philadelphia. The green and red in this map shows "working class" and "middle class." Source: William Edward Burghardt Du Bois and Isabel Eaton, The Philadelphia Negro: A Social Study, 14 (published for the University, 1899).



1.14 W. E. B. Du Bois's chart shows the average household budget of African Americans in Atlanta. Source: W. E. B. Du Bois, [The Georgia Negro] Income and Expenditure of 150 Negro Families in Atlanta, Ga., U.S.A., circa 1900, drawing: ink, watercolor, and photographic print, 710 x 560 mm (board), circa 1900, LOT 11931, no. 31 [P&P], Library of Congress Prints and Photographs Division, https://www.loc.gov/pictures/item/2013650354/.

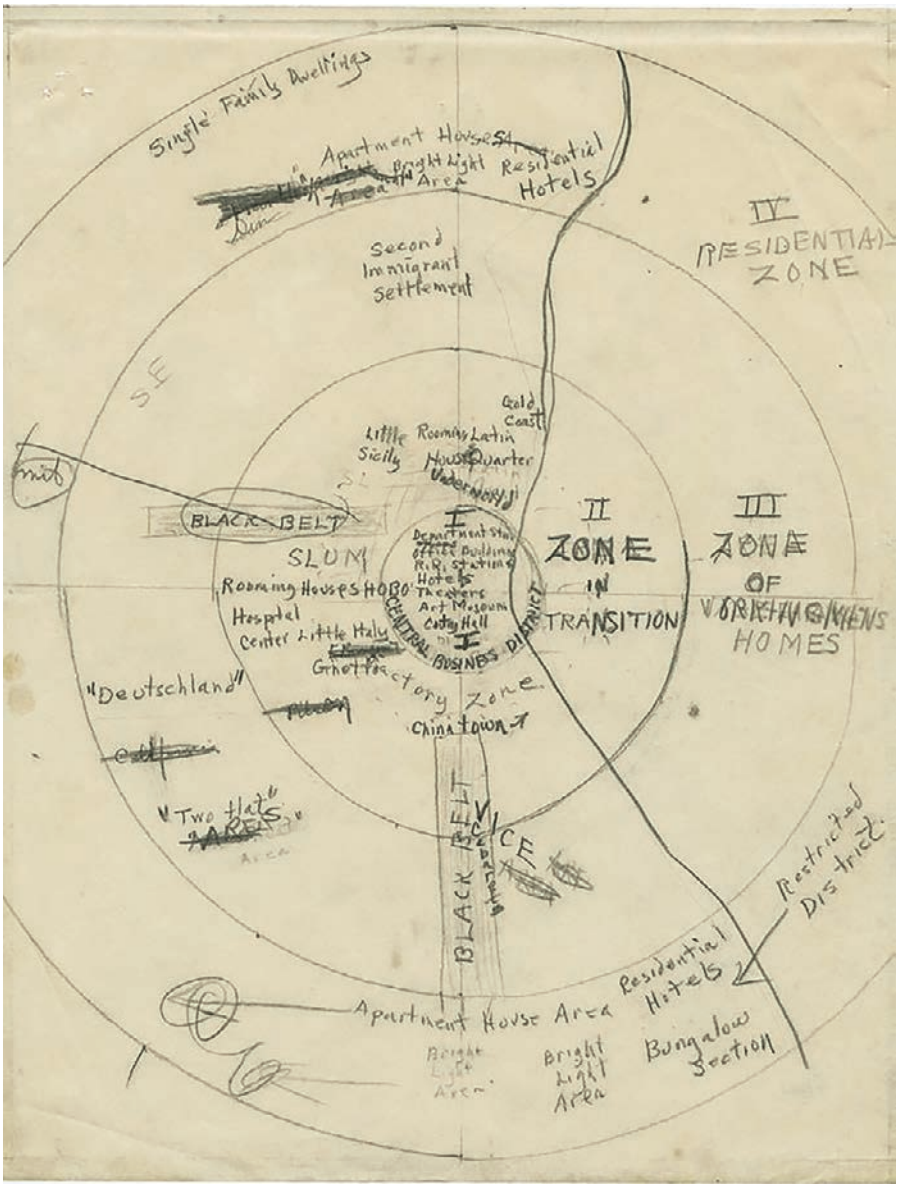


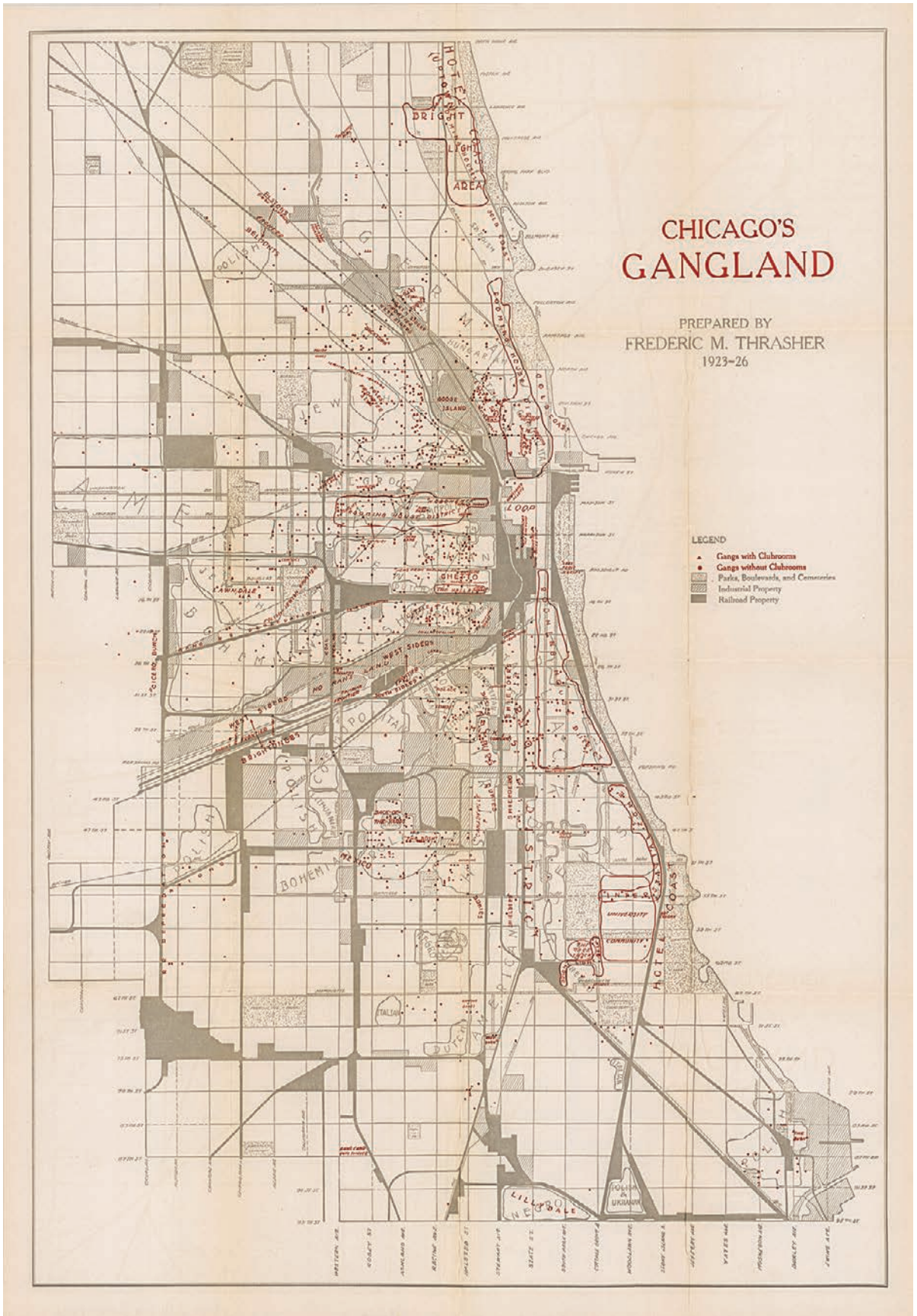


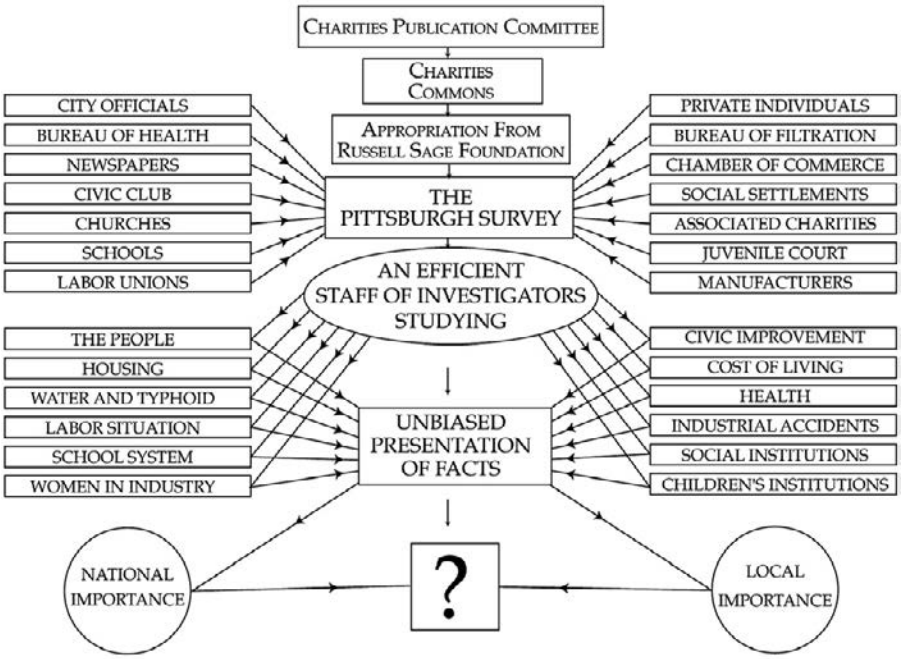
1.15 W. E. B. Du Bois's chart shows the types of businesses owned by African Americans in the United States, illustrating that black men had jobs similar to those of white men and helping to elevate the status of African Americans. Source: W. E. B. Du Bois, [A Series of Statistical Charts Illustrating the Condition of the Descendants of Former African Slaves Now in Residence in the United States of America] Negro Business Men in the United States, circa 1900, 1 drawing: ink and watercolor, 710 x 560 mm, circa 1900, LOT 11931, no. 57 (M) [P&P], Library of Congress Prints and Photographs Division, https://www.loc.gov/pictures/item/2014645363/.

1.16 The first hand drawn map of Burgess's famous theoretical model used to explain the social organization within urban areas. The diagram is central to the Chicago School of Sociology's work. Source: Ernest Watson Burgess, "Map of the Radial Expansion and the Five Urban Zones," n.d., Ernest Watson Burgess Papers, University of Chicago Library.

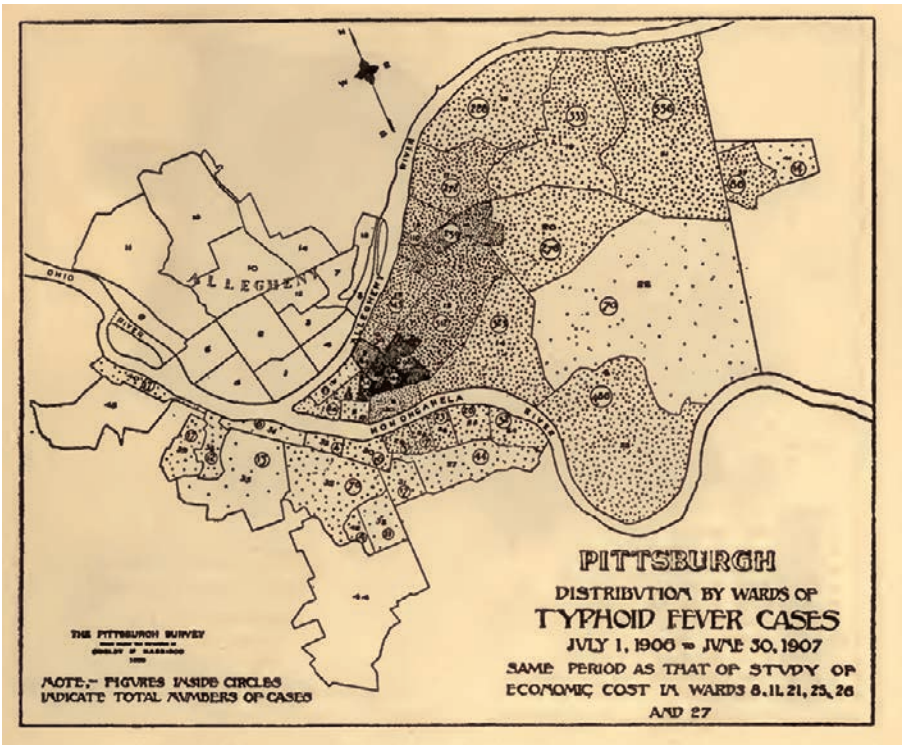
1.17 This map accompanies Frederic Thrasher's *The Gang: A Study of 1,313 Gangs in Chicago*, a text about the Chicago School of Sociology. This map was used to argue that it is not just geography that causes gang behavior but also the breakdown of institutions. Source: Frederic M. Thrasher, *The Gang: A Study of 1,313 Gangs in Chicago.* Chicago: University of Chicago Press, 1927.



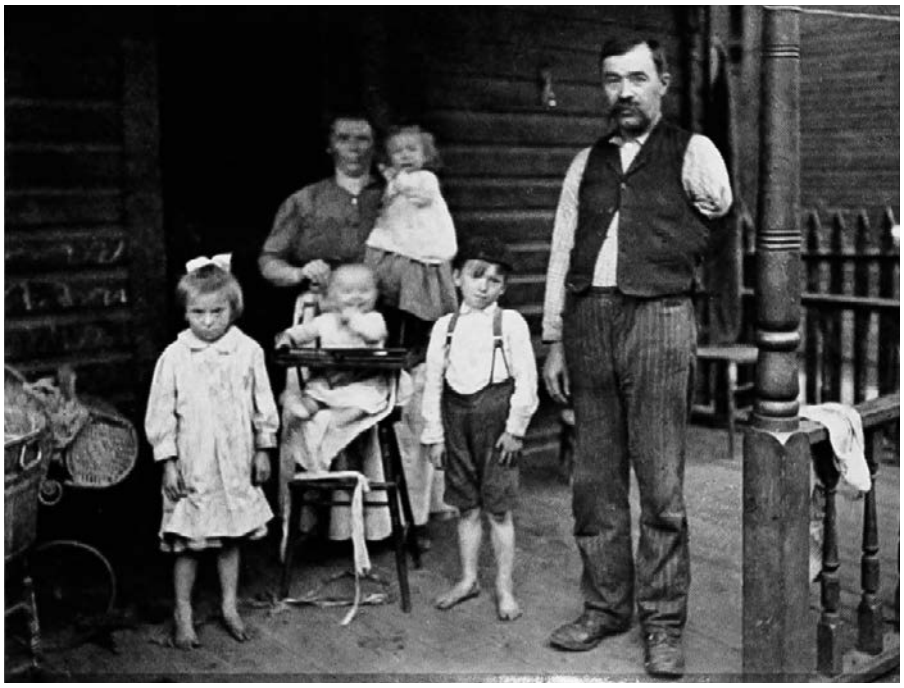




1.18 The authors of the Pittsburgh Survey developed multiple graphic devices to share their work. Source: Paul Underwood Kellogg, "Figure 1.1," in The Pittsburgh Survey: Findings in Six Volumes, ed. 1909 (112 East 64th Street, New York, NY 10065: © Russell Sage Foundation, n.d.).



1.19 Maps and photographs were an important part of the Pittsburgh Survey. Source: Paul Underwood Kellogg, The Pittsburgh Survey; Findings in Six Volumes, vol. 5, 1909, https://archive.org/details/pittsburghsurvey05kelluoft/page/xvi. Source: Paul Underwood Kellogg, "Figure 1.1," in The Pittsburgh Survey: Findings in Six Volumes, ed. 1909 (112 East 64th Street, New York, NY 10065: © Russell Sage Foundation, n.d.).



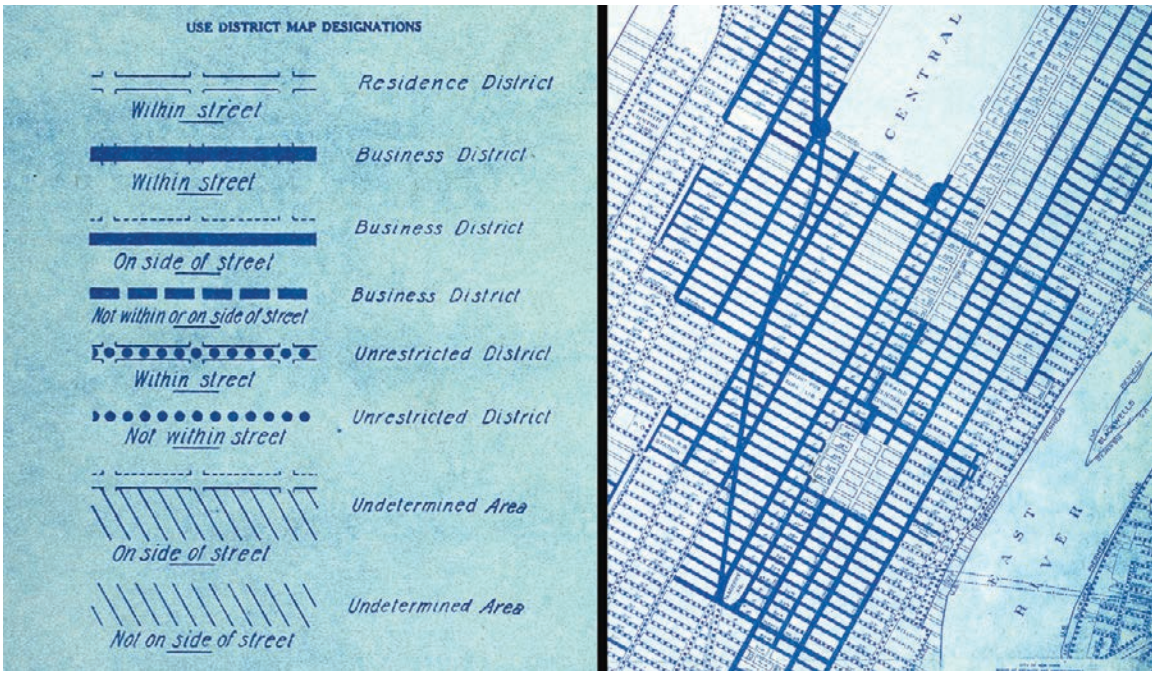
1.20a and 1.20b Photographs taken during the Pittsburgh Survey. Source: Margaret Frances Byington, Paul Underwood Kellogg, and Russell Sage Foundation, *Homestead: The Households of a Mill Town*, 1910, https://archive.org/details/homesteadhouseho00byinuoft/page/n189; Crystal Eastman, *Work-Accidents and the Law*, 1910, https://archive.org/details/cu31924019223035/page/n231.

they reinforce ideologies steeped in structural racism. What is perhaps more problematic is that even though the Chicago School learned its mapping techniques from Addams and DuBois, it marginalizes their connection to this earlier work by a woman and an African American sociologist.[51]

The social work being done to understand the poor in Chicago (by both Addams and the Chicago School of Sociology) influenced fields beyond sociology and was widely adopted by city planners. This work also influenced philanthropic organizations. The Russell Sage Foundation in particular funded extensive surveys on the social conditions in numerous US cities. Among the most notable was its *Pittsburgh Survey* in 1907 (figures 1.18 through 1.20b), which collected extensive data on living conditions, sanitation, light and air, education, and poverty, among other characteristics.[52] Hiring specialists from numerous fields, Russell Sage researchers made use of photographs, maps, and charts to communicate their results, and set the standard for others gathering qualitative and quantitative data to tell the story of the industrialized city. Data, both qualitative and quantitative, became crucial to changing the quality of life of people in cities everywhere.

Cadastral Mapping: Modern City Planners

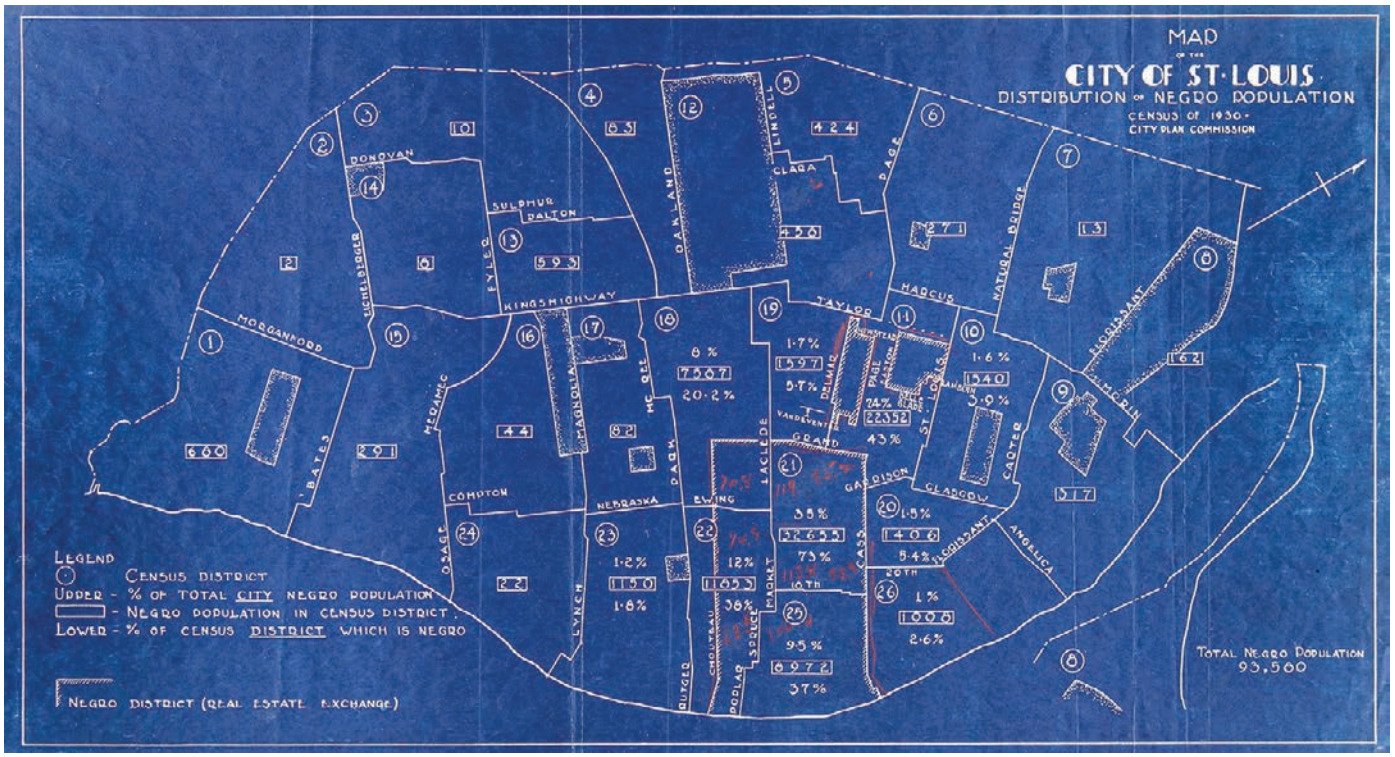
At the turn of the twentieth century city planning had yet to become a profession, so city planners were often trained as civil engineers. They took cues from the surveys performed by sociologists such as Addams, Du Bois, the Chicago School, and the Russell Sage Foundation to develop zoning and land-use plans that they hoped would harness the benefits of rapid urbanization. New York City's first zoning plan was established in 1916 and set out to separate incongruent uses. The new plan zoned Midtown West for manufacturing (largely of apparel) to provide a location separate from tenement housing. Early engineers, who were often referred to as technocratic planners, used detailed surveys to map out their cause, just as Addams and her colleagues did. Their objectives, however, often differed because they were largely hired to help increase the economic prowess of cities rather than address their social needs, such as improving poverty conditions. In his book *The Birth of City Planning in the United States*, Jon Peterson explains how this marks a shift from planning ideals that come from moral reform to those based on scientific progressivism.[53] Technocratic planning is based on a conception of decision making anchored in objective knowledge, such as



1.21 City of New York Board of Estimate and Apportionment, Use District Map, 1916. This is considered one of the first zoning maps. Source: https://digitalcollections.nypl.org/items/510d47e4-80df-a3d9-e040-e00a18064a99.

data, that will yield maximum benefits for the public.[54] It involves collecting extensive data and using it as evidence to develop zoning regulations to shape the city (see an example of zoning use designations in New York City in figure 1.21, considered one of the first zoning maps).

Harland Bartholomew professionalized this new scientific approach that introduced zoning as a main tool for land-use planning. He relied on data to develop his planning ideas, using it as evidence to support his own beliefs and ideologies—which, to put it mildly, often lacked concern for those on society's margins. For Bartholomew, the economic health of the city was the most vital concern. He believed removing the blighted areas of the city and improving access to its central business district would help the city's economy and increase its overall wealth, which in turn would be good for everyone.[55] This vision is reflected in his numerous plans. They emphasize slum removal, the development of highways to enhance efficiency over maintenance of



1.22 This 1934 map, based on the 1930 census, shows both the distribution of the black population and the boundaries of the St. Louis Real Estate Exchange's 1923 "unrestricted" zones. Source: City Plan Commission, "Map of the City of St. Louis: Distribution of Negro Population, Census of 1930 (Realtors' Red Line Map) 1934, http://collections.mohistory.org/resource/221591.

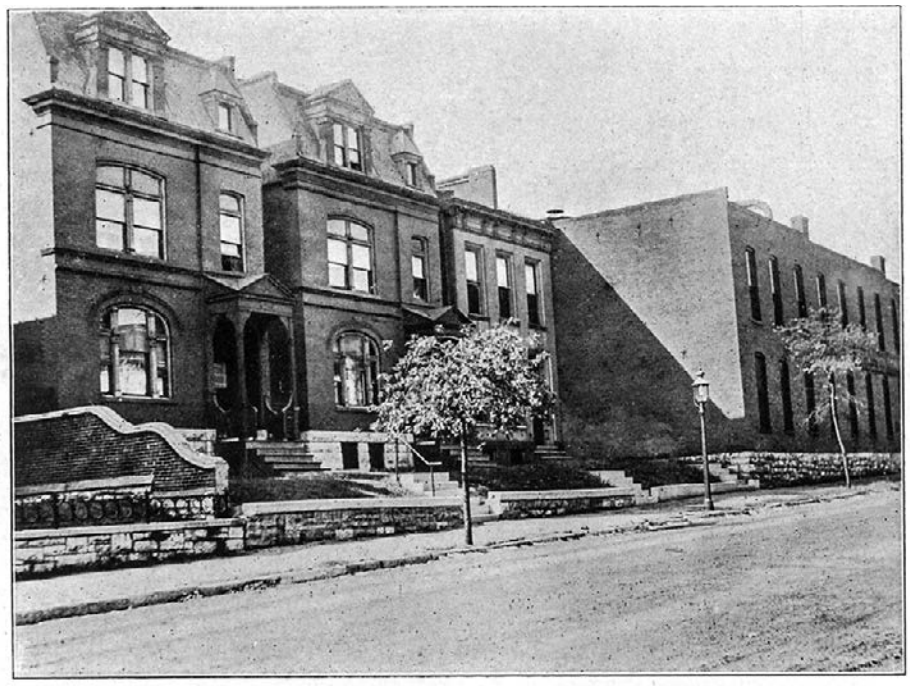
existing communities, and a focus on development in business districts in the core of the city. His analyses often specifically mention the need to remove or marginalize African American neighborhoods.

Bartholomew's approach, especially in the city of St. Louis, Missouri, illustrates the problematic ways in which city planners have employed data in the past, a legacy that helped facilitate the systematic segregation of US cities. Bartholomew became St. Louis's first city planner in 1915 and served in that capacity until 1950. He began his work there by implementing extensive citywide surveys of the character and use of buildings, using the Sanborn maps as his base. His surveyors visited each property and collected detailed information on its use and ownership—including the race of the owner(s) and/or occupant(s) of each residential building in the city.

Bartholomew used this data on race primarily to determine which neighborhoods were largely populated by African Americans, and then zoned those neighborhoods for undesirable uses, such as brothels, liquor stores, and

manufacturing, helping to encourage blight in once middle- and lower-class African American neighborhoods. Many found that the zoning caused their property values to significantly decrease; as a result, when African Americans wanted to move, they could not afford property in equivalent middle- or lower-class white neighborhoods (figure 1.22). In his own words, his goals were to preserve "the more desirable residential neighborhoods" and to prevent the movement into "finer residential districts . . . by colored people."[56]

Bartholomew used the data to develop the zoning ordinance of 1919, which slated areas in and adjacent to primarily "colored" communities for industry and manufacturing, making them less attractive to middle-class residents who were predominantly white.[57] The ordinance doesn't specifically mention race because zoning based on race was indeed illegal at this time, but the results of the new ordinance created a segregated city. African American communities were pushed to the margins. What is more problematic, using this plan as an example, the St. Louis zoning commission often rezoned neighborhoods for industrial use if it saw African Americans moving to those locations (figure 1.23).[58] This practice would make it hard for potential residents to receive mortgages, helping to further blight these communities.



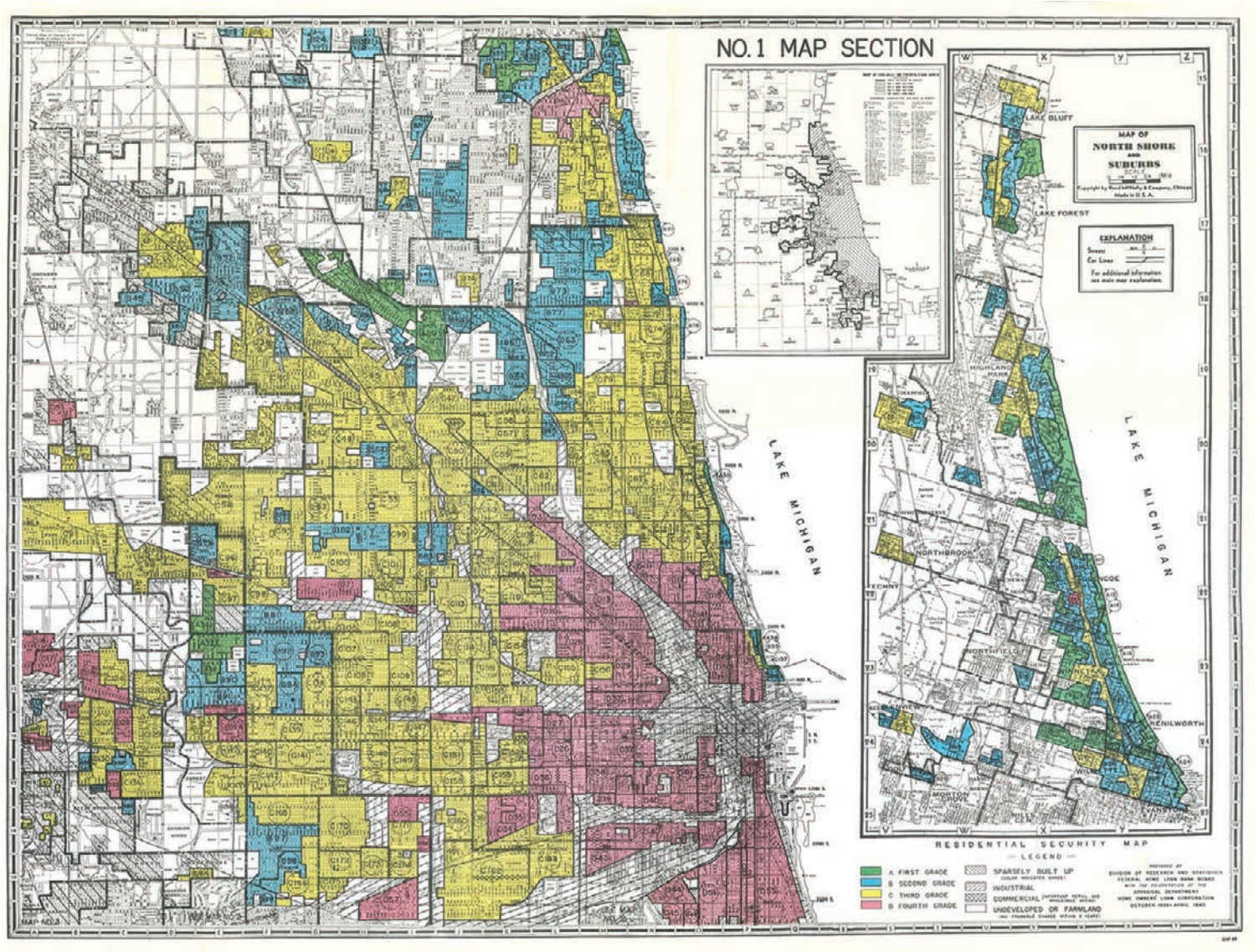
1.23 Bartholomew classified many African American neighborhoods in St. Louis as industrial districts. This photo shows a factory located right next to housing. Similar juxtapositions caused many of these neighborhoods to face severe decline. Source: Harland Bartholomew and City Plan Commission, Zoning for St. Louis: A Fundamental Part of the City Plan (St. Louis, Mo., 1918), https://ia902708.us.archive.org/5/items/ZoningForSTL/Zoning%20for%20STL.pdf.

The St. Louis zoning commission continued to use these exclusionary practices under Bartholomew's lengthy tenure as chief planner and beyond. By allowing land use that was not permitted in other residential neighborhoods—polluting industries, taverns, liquor stores, nightclubs, and brothels—the segregation helped turn these African American neighborhoods into slums.[59] African American homes in these neighborhoods were not eligible for Federal Housing Commission mortgages because their terms did not provide for incongruent use. Ironically, these zoning and use designations helped St. Louis garner urban renewal funds that paid for the ultimate clearing of many of these communities, which in one case led to the design and construction of the infamously dysfunctional Pruitt-Igoe housing project.[60] Some might argue that the clearing of these communities was always the goal of St. Louis's exclusionary zoning practices.

During the course of Bartholomew's career his firm developed over 563 comprehensive plans and his books and lectures influenced both the professionalization and the evolving profession of planning.[61] His methods, detailed in *Urban Land Uses* (1932, updated 1955), reached a broad audience, and the book became the standard text for academic planning programs across the country.[62] Using data resources amassed by Bartholomew's firm during the time it took to develop these numerous plans, the book provided a comparison of land-use and zoning practices (and their outcomes) in cities across the United States. Bartholomew also lectured at influential planning schools including Harvard and the University of Illinois, which helped to further circulate his ideas. The technocratic field of planning that Bartholomew helped to develop later became synonymous with the marginalization of the urban poor, giving rise to distinct practices of racial segregation.

Cadastral Mapping: Groundwork for Lasting Oppression

The types of cadastral maps and zoning ordinances developed by Bartholomew's firm laid the groundwork, literally and figuratively, for the practice of "redlining," or discriminatory lending practices based on the geographic location of properties. Community groups in Chicago's Austin neighborhood in the 1960s coined the term "redlining" in reference to the red lines that lenders and insurance providers admitted to drawing around the areas where they would not provide mortgages.[63] Bartholomew's zoning maps provided the impetus for disinvestment in African American communities that the Home Owners' Loan Corporation (HOLC) maps (1933–1936) continued (figure 1.24). HOLC maps identified mortgage risk in 239 cities



1.24 Home Owners' Loan Corporation maps of Chicago, showing the inner-city African American neighborhoods in red. Source: Form-Based Codes Institute, "Zoning for Equity: Raising All Boats," 2019, https://formbasedcodes.org/blog/zoning-equity-raising-boats/.

across the United States, using four categories. The highest risk—the red color on HOLC maps—identified African American communities, areas with so-called incongruent uses and poor housing stock.

The base data that HOLC used was developed from the 1934 and 1939 Works Progress Administration's Real Property Surveys, which mapped major cities across the United States, and which provided jobs during the Great Depression. The process of making the HOLC maps was overseen by the HOLC's Department of Economics and Statistics, which asked residents questions directed to identify risk, such as whether they held mortgages or felt they could hold one. Since Bartholomew was involved with the development of this program, many of the detailed, data-collecting strategies he used for his own work were used here too. According to Amy E. Hillier, these maps were given to mortgage lenders to identify areas where loans should not be given. The maps had a huge impact on helping to further marginalize these neighborhoods.[64]

Redlining maps disguise visible structural bias as objective truth. They also illustrate how the decision-making practices of banks were typically racially biased. These maps have become important symbols for how visualizations can be used powerfully, often by the powerful, to oppress the powerless. These maps permanently marked these urban landscapes as "other," places not worth investing in, and the ramifications of that labeling can still be seen today.

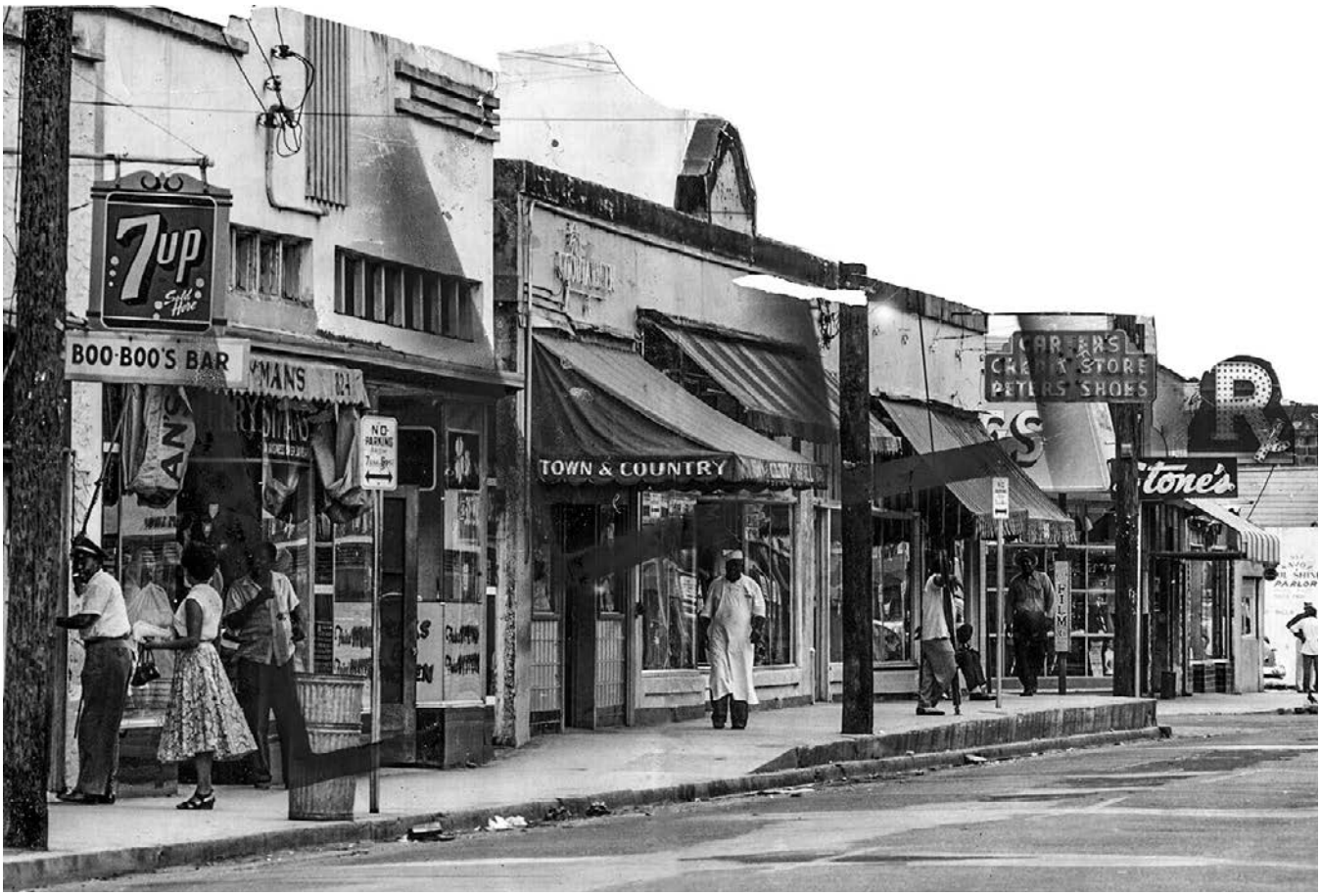
Data and Urban Renewal

Technocratic planners continued to apply data and mapping to urban renewal policies in the 1950s, securing funds from federal highway and housing grants (the 1949 Housing Act, for example) that could be gotten only if data were submitted as evidence in the application for funding. In cities across the country, technocratic planners devised ways to send highways through thriving neighborhoods they designated as "blighted" in the hopes of permanently removing these areas, using data as evidence. Perhaps one of the most notable occurrences of this phenomenon—the changes made to the original Interstate 95 highway plan for Miami—resulted in the total obliteration of Overtown, the city's historic African American neighborhood.

In 1955, Miami transportation planners set the course for the I-95 highway route to be run along an abandoned rail corridor, on the edge of a

residential area, and through "low-value" industrial tracts "to preserve and help protect existing residential neighborhoods and promote an economically desirable use of land."[65] But a year later, the highway route was changed when the state of Florida hired outside consultants—Wilbur Smith and Associates, a firm that had designed many other interstate routes across the country. The firm's engineers moved the route several blocks west to run through the inner core of Overtown, often called the Harlem of the South because of its lively music scene; they rationalized that the shift provided "ample room for future expansion of the [city's] central business district in a westerly direction."[66] The new plan had irreversible impacts on the city—the total removal of an African American cultural district.

Miami's business community did not see Overtown as a community that contributed to the economic viability of Miami (figure 1.25); they instead saw



1.25 This street scene of Overtown shows African American shoppers outside the retail establishments along NW 2nd Avenue between 8th and 9th Streets, circa 1950. Miami News Collection, HistoryMiami Museum, 1989-011-1703.

1.26 Massive highway interchange clears huge section of Overtown in 1967. Source: "Miami, FL: Overtown," January 3, 2019, http://miamiplanning.weebly.com/freeway-history--overtown.html.

an area that limited the city's economic growth, and they used data as evidence to make that point. Planning historians believe that the destruction of the neighborhood (figure 1.26) aligned with Miami businessmen's racial goals for their city, and according to several accounts, with their conviction that the African American presence would stifle the city's development.[67] This pattern of racist reasoning became standard practice during urban renewal and led to the destruction or marginalization of many more African American neighborhoods across the country, with data used as evidence to garner the necessary funds to make it possible.

Qualitative Analytics: The Call of Social Reformers

The failures of urban renewal policies, which often exacerbated problems in urban areas, caused some urban activists to fight the top-down planning methods that were being employed by technocratic planners. Jane Jacobs, an influential civic activist who became well known for her work in New York City in the 1960s, along with other civic actors such as Herbert J. Gans, led this charge.[68] Jacobs, who lived in New York's Greenwich Village, believed human-centered narratives were essential for understanding the economic and social needs of cities.[69] Jacobs fought against using data-driven policies as evidence for building highways because the results were so devastating to entire neighborhoods—and saw the removal of these neighborhoods as a complete failure to acknowledge the importance that their residents and networks played in the larger economy of the city.[70] Jacobs believed the public's voice should speak louder than data in the decision-making process.

It is interesting to note that neither Jacobs nor Gans argue against data *per se*, but rather against how data was typically collected and analyzed. Their focus was on understanding how the social connections of place created certain ties that strengthened the urban economy. Their interest leaned more toward qualitative data, including local stories and imagery. Unlike the data employed by the technocratic planner, this data included the perspective of the public because it was built by them.

Jane Jacobs's ideas about the city were often pitted against those of Robert Moses, New York City's most influential technocratic planner, who focused on making New York a "modern" city of highways and high-rises. Moses was able to deploy urban renewal strategies that effectively severed community ties, similar to what happened in Overtown. He did this using data as

1.27 "Mrs. Jane Jacobs, Chairman of the Comm. to Save the West Village, holds up documentary evidence at press conference at Lions Head Restaurant at Hudson & Charles Sts." Jacobs used here a different type of data—signatures of citizens. Source: Photo by Phil Stanziola, World Telegram & Sun, December 5, 1961, LC-USZ-62–137838, Library of Congress, https://www.loc.gov/pictures/item/2008677538/.

evidence. Jacobs, in contrast, came to represent a more qualitative planner who understood that the social networks of people are what make the city a great place to live, and that preserving those social networks, which help build and maintain community, were a legitimate focus of the urban planner's concern. For Jacobs, such social and support networks were often essential for low-income communities—and the modernist city as Moses envisioned posed a great threat to them.

Moses and Jacobs are often cited to represent two forms of city planning: the technocratic planner, who preferences the use of data to create evidence for modernist ideals, and the social planner, who preferences working to strengthen communities through building policies that support their needs and reinforce social ties. While the two are often juxtaposed, one using data and the other not, I would argue that both used data to create evidence for

1.28 Robert Moses, Park Commissioner of New York City, with a 1939 model of the proposed Battery Bridge. Moses is known for using data as evidence to send highways through poor communities. Source: Photo by C. M. Spieglitz, *World Telegram and Sun*, 1939, LC-USZ62-136079 Library of Congress, http://www.loc.gov/pictures/item/2006675178/.

their visions of the city, but they did so to different ends. In fact two famous pictures (figures 1.27 and 1.28) illustrate this eloquently: Jacobs is holding documents with citizens' signatures, while Moses stands in front of one of his urban renewal projects holding a report full of data analytics offering evidence for the development. Both are powerful documents, powerful pieces of data. Jacobs used her data to save a neighborhood from destruction, and Moses displaced whole communities to realize his modernist dream. Here we can see how using data for city planning can be a double-edged sword: it can be used to empower some and marginalize others.

With both sides armed with their own forms of data (qualitative versus quantitative), the debates between the technocratic planners and community advocates helped drive the development of public participation planning, which sought to narrow the gap between those who had control of information

and decisions and those who did not. Shelly Arnstein produced seminal work on this topic to advocate for greater knowledge sharing between what she called the "power-holders" and the "powerless" in public participation.[71] Inherent to Arnstein's approach was the need to engage broader publics, including communities typically marginalized by public participation processes, because they do not have the resources and data to advocate for their needs.[72] Interactive education and community dialog became important tools to fill that gap. Similarly, James Glass, who writes about citizen participation in planning, believed that information exchange that involved education, building support, supplemental decision-making, and representational input could change power relationships within public engagement strategies.[73] John Forester, a leader in the development of the field of participatory planning, believed that through interactive dialogue, power could be transferred, as dialogue allows groups to present their perspectives and ideas and thereby educate each other on critical issues.[74] This early work was influential in my development of the Data Action method. The idea that engagement with data is one way to help overcome the oppression it might cause.

As opposition to the methods of technocratic planning rose in the 1960s, the urban theorist Kevin Lynch began to develop methods for transforming the map from a tool of control to one of public expression. His method, detailed in his book *The Image of the City*, involved asking community members to draw "mental maps" of their cities.[75] Behind his cognitive mapping exercise is the idea that the elements of the city most important to community members will be those details they remember and document on maps they draw from memory. Therefore, looked at collectively, these cognitive maps show what is important to the people who live there.

Maps: Good or Bad?

What are we to make of these numerous examples that show how maps and data have been used to marginalized populations? Should we avoid maps as a tool? Avoid data? The title of this book might provide a clue. Data can be used for action: it's how we work with data that determines its potential to help society. Ultimately these historical tales show that maps and data reflect the biases of the people who are using them as a policy tool. Booth's maps had the consequence of removing whole communities from the city. Bartholomew's maps were used as a tool to systematically and purposely segregate cities.

Later we see the federal government using the HOLC maps to foster further disinvestment in communities. In each case, these maps could have aided these communities rather than push them down. Sociologists such as W. E. B. Du Bois attempted to use the same type of maps as a counter-narrative to the prevailing idea that African American communities were full of "unemployed criminals." Each map has its own history, and each cartographer had an agenda. They all used data to support their ambitions—aspirations that often left those on the margins of society without the ability to fight back against oppressive policies.

To use data for action, we must interrogate both our intentions and the intentions of others. That is why it's so important that we show the maps we make to the people whose lives are included in the data that these maps rely on. Then we must ask: Does our analysis ring true to them, or does it reflect our own biases about the communities we seek to help?

The Technology of War: Data and Cybernetics

After World War II, we see a great deal of military technology applied to everything—and that included city planning. Planners shifted from being inspired by sociology to being inspired by the efficiencies in management developed during the war. In *Warfare to Welfare: Defense Intellectuals and Urban Problems in Cold War America*, Jennifer Light discusses how experts at governmental think tanks and aerospace companies in the 1940s, 1950s, and early 1960s found new work applying their ideas to cities in the 1960s and 1970s.[76] Cities were looking for new strategies, and these specialists were looking for new ways to apply their expertise, especially as government funding for the military was starting to decrease. The federal government began to fund research institutions with an expertise in military technology such as RAND, SDC, and SAGE and asked them to apply their techniques toward civilian use.[77] The Johnson Administration's "Great Society" used modeling projects developed by the Pentagon and RAND to support this shift toward cities.

City planners turned to everything from cybernetics and computer simulations to satellite reconnaissance.[78] According to Peter Hall, writing in 1989 for the *Journal of the American Planning Association*, for city planners during this time, "anything that could not be expressed in numbers was inherently suspect."[79] Data was central to these movements because it

provided decisions with legitimacy, as the use of data was often equated with truth.[80] Los Angeles is a great example of a city that integrated some of these methodologies early on, developing a Community Analysis Bureau (established 1966) whose charter was to use data to make recommendations for everything from crime rates to unemployment and traffic.[81] Reports that came out of this bureau encouraged policies to increase the minimum wage, to help enroll children from the neediest families in preschools, and to establish affordable housing, among other things. But while the analysis proved useful, some argue that the work of the bureau was never truly integrated into the planning processes of Los Angeles; with no pathways to create planning action, it ultimately failed. Despite that failure, the bureau performed the data analytics necessary to secure important grant funding and turned to more policy-driven questions, unlike some of the models developed by enthusiasts of cybernetics.[82]

Cybernetics, championed at MIT by Norbert Wiener, is an interdisciplinary approach to examining how humans and machines communicate with and control each other. Its use for urban planning provided the ability to perform computational analysis that included theories of the city as a biological system and machine.[83] Jay Forrester, a researcher and professor at the MIT Sloan School of Management, applied cybernetics theory to modeling the dynamics of industry, including supply chains and resource flows.[84] After a chance meeting with John F. Collins, a former mayor of Boston, Forrester, who had no background in urban planning, set his sights on applying these ideas to cities, creating the Urban Systems Laboratory at MIT.

Forrester's experiment in modeling urban systems could be considered one of the biggest failures in using data analytics to understand, measure, and plan cities, and it helped to punctuate the problems of large-scale models to address planning problems. This failure stemmed from the fact that the model his lab developed, which was meant to be generalized to any city, made predictions that were completely out of line with the contemporary understanding of city planning. For example, it called for the clearance of slums to be replaced by luxury housing, and the inhabitants sent to public housing, which by 1969 had already proven to be an extremely problematic solution.[85] Some scholars attribute the failure of Forrester's model to its lack of focus on any one city—and therefore the variables could not be adjusted for any particular city—which made it hard to correlate the findings to real city behaviors.[86] Others noted the lack of understanding of the spatial dynamics

of cities and the interdependence of the inner and outer city (suburbs).[87] Perhaps the mostly highly cited critique, *Requiem for Large Scale Models* (1973) by Douglass B. Lee Jr., argues that large-scale models, such as Forrester's, don't start with theory-based policy questions, thereby creating "empty-headed empiricism."[88] Ultimately Lee calls on urban modelers to make models more transparent and focused on specific policy questions rather than dealing with the city as an abstraction.

Inaccurate models, high costs, and the reinforcement of hierarchical power structures ultimately led many cities to give up urban modeling developed out of the military in the 1970s.[89] The critical-planning theorist John Forester (not to be confused with Jay Forrester) argued that little consideration was given to the fact that data interpretations were heavily biased as a result of bias from those generating and collecting the information.[90] The construction of models was often flawed because of the lack of appropriate data to develop them, the use of biased information, and the fact of not testing whether the data results fit people's experiences. A telling example of this is the 1970s-era model developed by the RAND Corporation for the New York City Fire Department to determine the efficiency of its network of firehouses.[91] Basing the model almost solely on response-time data, the analysis led to recommending the removal of several fire stations in the South Bronx. The data model did not consider additional data factors such as gridlock, politics, multiple simultaneous fires (and therefore the need for back-up units), and the socioeconomic status of the neighborhoods in question. This resulted in an overload on demands for the fire stations that remained in the Bronx, and it is estimated that fires displaced half a million people during this period. The RAND model did not *cause* the rampant fires in the Bronx during the 1970s, but it did incorrectly assess how the closures would affect the other firehouses in the network, thus placing undo stress on the system and displacing many people.

An Urban Planning Tool: Public Participation GIS

The advent of desktop computing in the 1980s and early 1990s meant that data analytics of the 1960s were more accessible to the public. This ultimately led to the development of geographic information systems (GIS), often referred to as visual databases, which were designed with the capability to capture, store, and link geographic data and allow users to perform interactive queries and manipulate the data to extract new insights. Although the GIS

systems allowed for modeling, they also became a tool for managing city infrastructure—including detailed surveys of the cities performed by tax assessors to determine property values. The ill-fated models of the late 1960s and 1970s weighed heavily on the minds of early adopters of the software, who recognized that GIS could both marginalize and empower different populations depending on how it was used.[92] This led to a series of debates that began in November 1993, often referred to as the Friday Harbor meetings because of their location in the San Juan Islands of Washington State. During these debates, leaders in the GIS community discussed how it was difficult to use analysis from GIS to address social agendas because GIS software was often operated by those in a position of power.[93]

Ultimately these discussions led to the development of the field of Public Participation Geographic Information Systems (PPGIS), which attempts to address the power dynamics involved in spatial analysis and mapping by providing a methodology with which the public can more easily engage and learn about the data used to make decisions with GIS.[94] PPGIS projects are based on the premise that such engagement in data analysis and collection efforts can allow community groups to have more power in the political decision-making process.[95]

Since the development of PPGIS, governments and nongovernmental organizations (NGOs) have used it to share data with wider audiences and thereby influence national-level policies.[96] Community-based organizations have also used GIS within their own organizations to develop maps that represent their own perspectives and voices as they attempt to influence neighborhood-level planning initiatives.[97] Advocates for PPGIS believe these methods empower citizens by creating a process that makes it possible for the public to work with GIS systems and thereby critique the output, but still ensure that community interests are represented. GIS systems have now become an essential tool for planners, used for everything from evaluating environmental risk to managing electricity lines. Many open-source GIS software packages (including web-based mapping) are now available to truly put the power of modeling at the fingertips of urban planners as well as novice users.

The explosion of this software has created a new group of data enthusiasts who are not always aware of the risks of using data. This has produced a lot of "bad" or inaccurate analysis, not necessarily out of malice, but rather a lack of data literacy. For example, some maps floating around on the web

often misrepresent basic census demographics, as data novices don't know that they have to normalize census data for population and areas. More problematic, because the work is represented as a map, is the fact that the general public often does not question its legitimacy, as maps are so often equated with truth. Data literacy is essential for these data enthusiasts as well as the general public.

Smart Cities and Big Data: Urban Modeling Rebooted

In the last decade there has been an upsurge of interest among private and public agencies in the uses of data for the design, planning, and management of cities and urban environments. Excitement around the potential of "big data" to change the way we see the world has created an enthusiasm for applying data to civic action and policy change. IBM developed its "Smart Cities" and "Smarter Planet" advertising campaigns in 2008 to promote the use of technology and data to analyze the problems of cities.[98] In 2010, IBM created a partnership with Rio de Janeiro to make it a model Smart City.[99] During that same year, Cisco launched its "Smart and Connected Communities" program, which is based on using data analysis and web-based interface programs to connect cities through technology.[100] Also in 2010, the *Economist* magazine developed a series of stories called the "Data Deluge," which explained the possibilities of using data to develop strategies for almost anything, including cities. The hype around these projects generated fascination in the popular press and media for what the analysis of big data can offer cities.[101] Microsoft launched its CityNext program, intending to draw from its worldwide network of technology experts to make cities better places.[102] In 2015, Google launched Sidewalk Labs, a grouping of "urban innovation organizations" working on various projects including data analytics with the US Department of Transportation and developing a "smart city" solution in Toronto.

Stephen Goldsmith and Susan Crawford document this interest in using data to drive government policies in their book *The Responsive City: Engaging Communities through Data-Smart Governance*, which illustrates how governments are harnessing the power of big data to increase efficiency in the delivery of city services. In the foreword, Michael Bloomberg, the former mayor of New York, explains why he thinks data is essential: "If you can't measure it, you can't manage it." He goes on to say that, "Harnessing and understanding data helped us to decide how to allocate resources more

efficiently and effectively, which allowed us to improve the delivery of services from protecting children to fighting crime."[103] For Bloomberg, data analytics are essential for proper governance.

Urban planners have begun to shed their fear of data, which developed during the late 1960s largely in response to technocratic planning. In the fourth (and, to date, most recent) edition of *Introduction to Readings in Planning Theory,* Susan Feinstein and Scott Campbell insist that, like it or not, planners need "to deal with the coming flood of data." "Planners," they say, "need a larger conceptual world view to understand the ramifications of the digital revolution of the Internet, massive data storage and retrieval, and geographic information systems (GIS)." They argue that the questions planners have raised about reliance on "scarce, incomplete data" in the past may have been supplanted because data is now richer, and the processes to analyze and visualize it have been more fully developed.[104] Many planners embrace data, but we should continue to regard the intent of data analysts with a healthy dose of skepticism.

Open-source data access and online technology tools have allowed many other people, from opposition activists to NGOs to ordinary citizens—to create their own maps as well. GIS and computer mapping software are increasingly available, as are web-based mapping tools, allowing all sorts of non-traditional data visualizers to benefit from the potential these technologies offer to expose spatial patterns. The data used in these systems is also becoming more accessible, as cities and private data producers realize that opening data to the public means opening up possibilities—again, hopefully—for improving the cities, and for allowing everyone the chance to build upon and analyze the knowledge gained through public data collection. However, just because the data is more readily accessible doesn't mean everyone has the skills and knowledge needed to make use of it.

The misguided models of the 1960s and 1970s have largely been forgotten and replaced by an excitement for harnessing the power of big data for cities. Large corporations such as IBM, Cisco, and Microsoft, which can make a profit by applying their computing power to big data analytics in cities. Yet these analysts, who are often untrained in urban planning or related fields, are not always familiar with how data models have been used inappropriately in the past. It is therefore up to us to educate them to use data ethically and responsibly.