TYPEFACE: MACHINE-VIEWING GENTRIFICATION ON STOREFRONT IMAGERY IN BEDFORD-STUYVESANT, BROOKLYN

by

ALEXANDER MCQUILKIN

A master's capstone project submitted to the Graduate Faculty in Data Analysis and Visualization in partial fulfillment of the requirements for the degree of Master of Science, The City University of New York

2026

APPROVAL

Typeface: Machine-Viewing Gentrification on Storefront Imagery in Bedford-Stuyvesant, Brooklyn

by

Alexander McQuilkin

This manuscript has been read and accepted for the Graduate Faculty in Data Analysis and Visualization in satisfaction of the capstone project requirement for the degree of Master of Science.

Approved: January 2026

Matthew Gold, Advisor

Matthew Gold, Program Director

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

Typeface: Machine-Viewing Gentrification on Storefront Imagery in Bedford-Stuyvesant, Brooklyn

by

Alexander McQuilkin

Advisor: Matthew Gold

Gentrification—broadly, the replacement of a less powerful group by a more powerful one in an urban context—is oft-discussed in the popular press, but is ill-defined or ignored in sociological and economic academic discussions. Furthermore, academic treatments of displacement understandably focus on measurable yet fairly abstract indicators like changes in rent or income, whereas neighborhood change is often registered by residents on the ground using visual, but difficult-to-quantify markers like retail turnover. This project uses image recognition technology on a set of storefront photos to index the visual streetscape of a neighborhood, as well as to track changes to that portrait over time, and presents the findings in an accessible format on the web. The model relies on a binary classification developed by New York City cultural anthropologist Edward Snajdr and sociolinguist Shonna Trinch (2020), wherein colorful, text-heavy "old-school" storefront signage evokes openness, diversity and accessibility, while the spartan, symbolic and glass-laden "new-school" signals clubbiness, cultural capital and upscale. The neighborhood focus is Bedford-Stuyvesant, a historically Black section of Brooklyn that has lately been in the top three community districts for proportional increases in rent, income, and white share of the population. The model finds a small but significant likelihood that old-school stores have closed, while more newly opened stores reflect the new-school style.

ACKNOWLEDGMENTS

I would like to thank program director Matthew K. Gold for his trusting and encouraging advisement throughout the process, Stefano Morello for taking the time to listen and offer resources and direction, instructors Omar Nema and Deborah Balk for helping shape the project's earlier iterations and offering invaluable technical upskilling, Jason Greenwald, Chris Bradicich and the rest of the Live XYZ team for providing the data and believing in the project’s value, Shonna Trinch and Edwdard Snajdr for laying the groundwork to make the project possible and offering to potentially collaborate in future, and Emily Osgood, my fearless boss at HPD who has always believed in the synergistic opportunities of my pursuing this program.

CONTENTS

TABLES

FIGURES

DIGITAL MANIFEST

I. **Capstone Whitepaper (PDF)**

II. **Website**

a. Public site: https://alexmcqw.github.io/Bed-Stuy-signs/

b. Code: https://github.com/alexmcqw/Bed-Stuy-signs as well as in a .ZIP file (Bed-Stuy-signs-repo.zip)

c. Archival file (WARC format)

III. **Model**

a. Hosted on Teachable Machine: https://teachablemachine.withgoogle.com/models/lTysYEPbC/

b. Archival file (Bed-Stuy storefronts.zip)

IV. **Data**

a. LiveXYZ data download plus new fields (described in the Appendix: Data Dictionary): BKCB3export(Closed_operating_only) with predictions_updated.csv

b. Training data (hosted online: https://docs.google.com/spreadsheets/d/1PSPU3ZG-4tdYZpHq_FaGankImTOYDALRK2b2FgHT6gU/edit?usp=sharing) and file: Training data.csv

A NOTE ON TECHNICAL SPECIFICATIONS

This project is composed principally of a website (including HTML, CSS, and Javascript files), all of which are accessible in the GitHub repository: https://github.com/alexmcqw/Bed-Stuy-signs. The repository also contains the Python scripts used to batch-process the images which were stored locally. Links to all images can be found in the .CSV data file.

The image processing model is public and open-source, hosted on Google's Teachable Machine: https://teachablemachine.withgoogle.com/models/lTysYEPbC/ and also provided in the .ZIP file. It can be opened in a TensorFlow.js library.

"You are a New Yorker when what was there before is more real and solid than what is there now." - Colson Whitehead

"Photography is time travel, and the past surrounds us even as it's absorbed." - Chris Stein

INTRODUCTION & RESEARCH QUESTION

In 2025, the New York Sign Museum, which has been around since 2019, started getting so many visitors that it moved from an appointment-based visitation system to an online scheduler with thrice-weekly spots. Perhaps it was thanks to a write-up in the *New York Times*, which quoted the founder, David Barnett: "we're pushing back against all of these forces in the world — the accelerating globalization and homogenization of all parts of culture that would want to make every city in the world look the same...push back against that a little bit and really try to reclaim that visual language of place" (Kode, 2025).

Perhaps also, its message of preserving an eroding physical past from our cities resonates ever more in an era marked by frenetic technological change and visual culture coopted by AI sameness. My own trip to the museum came over summer 2025, partly out of personal curiosity and partly in preparation for this project.

One of my favorite takeaways, however, wasn't inside, but in the museum pamphlet. It traces the history of the East New York building (which also houses a sign-installation shop also owned by Barnett) the museum calls home through at least three other historical businesses, including a defense supplier and two clothing manufacturers, by revealing a palimpsest of signage hidden around the facade and inside (Barnett 2025).

Figure 1: Inside the N.Y. Sign Museum (Alexander McQuilkin)

That small gesture, of being curious about and acknowledging what came before, in a city as vast and old as New York, feels particularly important now, as we spend more and more time online and take for granted the way capitalism infuses the ways we move about and interact with our city. There's a meme that I frequently come back to, that encapsulates the silliness and topsy-turvy-ness of the tight interweaving of capital, power and culture that defines how we experience the urban condition in the 21st century. It's an image macro showing hipster haircuts, with the text superimposed: "If this haircut shows up in your neighborhood, your rent is going up" (https://www.instagram.com/p/DJuEpiHAgMd/, Hipsters of NY, 2025).

The concept of gentrification is actually so well-ingrained in internet culture that it's taken on definitions outside of the urban streetscape, to basically include anything upscaled, or really, white (Urban Dictionary, n.d.). But despite its sometimes lighthearted deployment, it's a serious topic that deserves continued scholarly treatment, particularly if "gentrification" itself—and not just its proxies like increases in incomes or rents—can be measured. This project attempts to fill part of that gap.

Visual indicators of upscaling of the neighborhood are not new. Bogeymen for gentrification have included horizontal fencing (Coleman, 2016), sans serif metal house numbers (Garza, 2023) and grey paint (Lang & Harden, 2025). Here, I have chosen a marker on which I have access to much photographic data—storefront signage aesthetics—and will try to tease out changes in that trend across time and space in one New York City neighborhood.

Retail turnover—both of what kinds of goods and services are on offer as well as how they're presented—is a visual, even emotional message about change around you, regardless of whether you think your neighborhood is "gentrifying," or if you don't closely track other, more mundane but quantifiable changes like real estate prices. These are the semi-public spaces where we spend time and money, and hone our identities, so the loss of a favorite restaurant or the arrival of an international chain can be unnerving.

I have noticed and thought about a dichotomy of retail storefront aesthetics for about as long as I've been interested in urban economics and equity. Older, more traditional vendors, these days often run by immigrants, often deploy a sincere, straightforward approach to self-promotion on their retail facades, something I would call high-information/low-design. Newer, maybe younger, and more media-savvy *entrepreneurs* opt instead for a style that's heavy in visual codes, winking "if you know, you know." This is low-information/high-design.

Conveniently, in the course of researching this project, I found a book that recognized a similar binary, and employed a methodical anthropological and sociolinguistic analysis of this phenomenon, also in New York City: Shonna Trinch and Edward Snajdr's *What the Signs Say: Language, Gentrification, and Place-Making in Brooklyn* (2020).

At the same time, image recognition technology has advanced to a degree that it's possible to do large-scale visual analyses of the cityscape. AllText.nyc uses text-identification on Google Streetview photos to establish a searchable index of the city's public textscape (Daniels & Zhao, 2025). While Google Streetview provides a sort of panopticon view of the city's streets (and necessitates some pretty advanced scraping and processing technology), my project is concerned with a much narrower component of the cityscape: the storefront. Live XYZ provides just that (more in the 'Data source' section below).

I was able to start familiarizing myself with the Live XYZ data in a project for an elective sociological geographic information systems (GIS) class here at the Graduate Center (SOC 81900: Spatial Demography), and the findings and approach of that project would go on to inform this investigation. For that project, I focused on retail category turnover in the three Community Districts (Brooklyn 2, 3 and 4) that showed the greatest increase in white share of the population, the greatest increase in high-earner share of the income distribution, and the greatest increases in rent over the last 20-25 years (Figure 2).

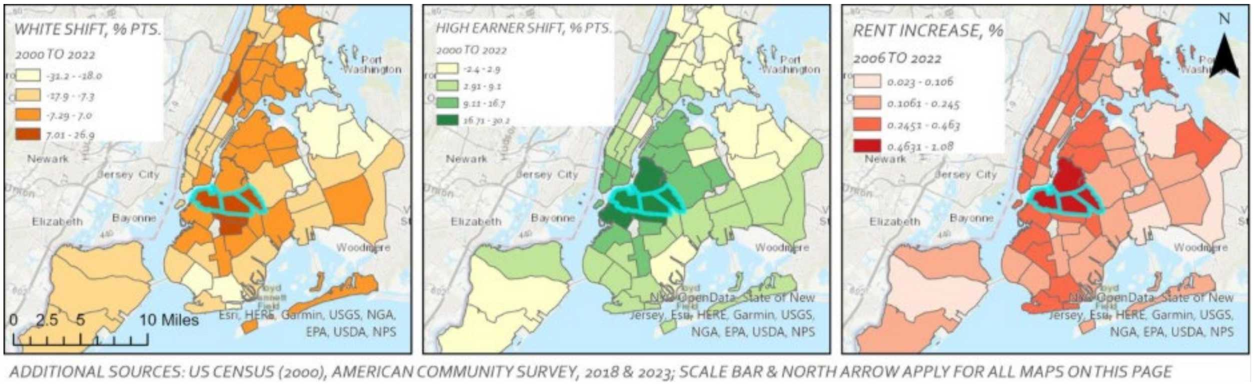


Figure 2: White shift, high earner shift and rent increase, varying timescales, with the three Community Districts (CDs) that scored high on all three outlined in teal

I called these three the "gentrification core" (Bedford-Stuyvesant is more or less contiguous with Brooklyn Community District 3, the middle of the highlighted districts) and found that between 2018 and 2025, this area saw a greater share of some categories of new storefronts than the city as a whole (thrift store +0.58 percentage points, tattoo parlor +0.44 pts., cocktail bar +0.43, coworking space +0.41, real estate agency +0.38, art gallery +0.36, and fitness studio +0.26) and a greater share of *other* categories of storefronts *close* compared to the city overall (churches +0.3 pts. and daycare centers +0.28) (McQuilkin, 2025).

Since the 1990s, Bedford-Stuyvesant (Bed-Stuy) has been the largest African-American neighborhood in New York City (Murray & Murray, 2023). But its Black history goes back much farther. In the 1830s, freedman James Weeks established a free Black community called Weeksville, in what is now considered Crown Heights, just south of Bed-Stuy. The area was heavily settled by African-Americans fleeing the Jim Crow South in the Great Migration to northern cities. Starting in 1938, the neighborhood was red-lined in the federal Home Owners' Loan Corporation (HOLC) maps, along with other immigrant and minority areas like the South Bronx, Lower East Side, and Harlem. (Moss, 2017)

Around the same time, Bed-Stuy earned the nickname “Little Harlem” as it became known as a locus of Black culture (Moss, 2017). In the 1950s the area fell prey to the common tactic of “block-busting,” wherein real estate speculators encouraged white flight to the suburbs and sold homes to Black residents at inflated prices, leading to overcrowding and other social stresses as new homeowners struggled to afford their mortgages (Moss, 2017). In the decade to 2010, the white population soared 633% marking the biggest change of any racial group of any New York City neighborhood (Moss, 2017). In recent years, Bed-Stuy has been at the heart of a trend of investors predatorily obtaining deeds to brownstones from long-time elderly Black homeowners or their heirs (Bradley-Smith, 2024).

It’s in this context that I pose the question: can we use public commercial aesthetics as a proxy to illuminate one facet of the contested capitalist city, and what are the benefits and drawbacks to leaning on machine learning technology in that pursuit?

LITERATURE REVIEW

What is 'gentrification'?

The term was coined by British sociologist Ruth Glass regarding a phenomenon she observed in London in the 1960s: "members of an upper class invade a lower-class neighborhood, purchase and upscale the houses, displace the people, and change the character" (Moss, 2017, p. 34). It is at root a power imbalance among classes, which is frequently intersectional with race and other identities (Moss, 2017).

In urban planning literature, the term is most often used in the context of housing insecurity, and is—misleadingly—lumped in with the phenomenon of displacement, wherein residents who were once stable in their homes, lose them, due to increasing land values and building sales, landlord harassment, government redevelopment schemes or some combination. But city and regional planning professor Lance Freeman (2006) and others (Dawid, 2014; Cortright, 2025) have found little impact on the wellbeing of existing residents when newcomers start arriving, and far greater harms to livelihood for those living in *unchanging* pockets of concentrated urban poverty.

Freeman even finds residents' anxiety about rising prices balanced out by excitement about new retail amenities (Chideya, 2006). Retail turnover is often the first or most visible indicator of neighborhood change, a transformation that plays out over decades. "One person's energizing excitement about new creations can be another person's concern about the existing neighborhood coming undone" (Trinch & Snajdr, 2020, p. 177).

Food and drink establishments—perhaps because eating and drinking traditions are so culturally imbued—tend to set off alarm bells for those who are paying attention (Cheung, 2017; Reid, 2020). Density of food establishment type (grocery store, *bodega*, farmers market) was found to be correlated with race in three U.S. cities, with negative health implications for racial minorities (Moore & Diez Roux, 2006). But the clear majority of the retailers discussed in this project are small businesses, not national chains (more on that in the 'Data collection' section below), and so it's important to remember the larger forces at play in discussing neighborhood change in New York City and elsewhere.

In 1979, CUNY Graduate Center anthropology and geography professor Neil Smith postulated that a "rent gap," or the gap between current rent as land is presently used and its potential, meant that in a neoliberal context, gentrification becomes an intentional process not spearheaded by individuals, but organized by government and corporations to exploit redevelopment and changes in land value (Trinch & Snajdr, 2020).

Smith also called Ruth Glass's London gentrification merely a first wave; a second one came in New York City with the 1970s fiscal crisis, and a third wave as "gentrification generalized," the current moment beginning in the 1990s and moving out from Manhattan: "new housing, new restaurants, new artistic venues, new entertainment locales...aimed at a social class quite different from those who populated" New York City neighborhoods prior to Mayor Bloomberg (in office 2002-13) (Moss, 2017, p. 39).

As mayor, Michael Bloomberg even penned an op-ed in the *Financial Times*, advocating the role of government and real estate interests to *pave the way for the individuals* (like artists and queer people) who would often then take the blame for gentrification: "cities must attract creative talent in order to compete for...tourists and businesses," e.g. with bike lanes, art galleries, etc. (Moss, 2017, p. 324).

Gentrification is a long game, and as neighborhoods move from poor to expensive along a long trajectory, they inevitably experience interim periods where it looks like they could go either way, with old-fashioned and newcomer indicators sitting side-by-side (Trinch & Snajdr, 2020). This will be evident in my own analysis that follows. Between 2004 and 2014 average commercial rent in the city increased by 34% (Moss, 2017). And starting in 2020, the Covid-19 pandemic threw up a host of new challenges to anyone operating or hoping to operate an in-person shop or restaurant.

'What the signs say'

I came across Shonna Trinch and Edward Snajdr's *What the Signs Say: Language, Gentrification, and Place-Making in Brooklyn* (2020) while browsing literature for this project. Not only was it a deep source of background literature (as I've been citing generously), it also confirmed in my mind the storefront style binary I had been observing myself, and would serve to inform the aesthetic structure I would build my model on. For Trinch, a sociolinguist, and Snajdr, an anthropologist, storefront signs are not just representative of Brooklyn's diversity and commercial vitality, but are totems of culture (2020).

As part of their project—which has continued even after the book's publishing—Trinch and Snajdr analyzed over 2,000 signs in 14 different neighborhoods across the borough, with an eye mostly to language (what businesses call themselves and how they advertise their wares in text), but also to aesthetics – how busy is the text on a storefront, how color and photography are deployed, the materials and cost of awnings and exterior decor, the use of supplementary signs and presence of products visible either through the window or directly on the sidewalk. Most small street-facing businesses, they found, fall into two camps: what they would term "old-school/vernacular" and "new-school" (Trinch & Snajdr, 2020).

This basically confirmed the pre-gentrification and gentrify*ing* styles that I had observed myself. Characteristics of the old-school signs included: non-standard spelling and syntax, multiple words and graphics and languages other than English, additional/ancillary signage and use of language "literally and explicitly" (Trinch & Snajdr, 2020, p. 55). They also employ clear, demonstrative offerings of what's being sold, and blatant use of sign size and color to draw attention.

The effect this had, from not only the authors' perspectives but from interviews of other professionals, students and everyday passersby, was to "make everyone feel welcome... straightforward... open to everyone... what you see is what you get" (Trinch & Snajdr, 2020, p. 56), a "directness...does not discriminate" (p. 57), an "aggressively democratic and tolerant system of commerce" (p. 58), a "collective sincerity...by removing barriers to communication, signifies inclusion and openness...democracy in diversity" (p. 63), "authentic, quaint, interesting" (p. 164), has the ability to adapt, "generative inclusivity" (p. 205), but also "lawless" (p. 163), "poor, shady, ethnic, outer-borough" (p. 164).

There was a completely different look that began to proliferate around the borough however, led by a new class of enterprising Brooklynite. This "new-school" was "textually very sparse...cryptic," with names incorporating multiple meanings and witty wordplay (Trinch & Snajdr, 2020, p. 46). The aesthetic was much more stripped down and modern, signaling "expertise," "elegant," sophisticated" (p. 63). Using much less text, these businesses broadcast a lifestyle using design and symbology, while not clearly advertising what goods/services are on offer.

Not only is the ambiguity *part* of the message, but perhaps these are destinations sought out on social media. A younger breed of entrepreneur might make diligent use of advertising and design not only in situ, but online as well. To observers, Trinch & Snajdr found these establishments signify "mystery/aspiration/clubbiness" (p. 63), "demure" (p.

228), "erudite and playful references to history, literature, and culture" (p. 67). The veil of in-the-know-ness is multifunctional: for a certain kind of upscale consumer, symbols like sans serif fonts and all-glass facades mark a welcoming beacon, while turning away passersby who might be confused by the dearth of helpful information. "Like a secret club" (p. 71), heard Trinch & Snadjr, an "in-your-face performance of...wealth" (p. 94).

It's interesting to note here how the City's observation of and guidance on storefront signage has changed over the years, perhaps in response to conversations around gentrification and power. Under Mayor Bloomberg, the Department of Small Business Services published a brochure on storefront improvement advice, decreeing, "the guiding principle of good storefront design is to keep it simple" (2012, p. 4) and "don't put too much information on signs. They appear cluttered and are hard to read" (2012, p. 13).

By 2019 however, the City Council passed an amendment to the Administrative Code—Local Law 28—that tasked the Department of Buildings with creating:

- A technical assistance program to help small business applicants seek and secure relevant permits for sign installation and educate them about signs and regulations and how they can move to legality;
- A temporary waiver of 75% of fees connected with sign installation permits; and
- An interagency task force to explore issues related to sign regulations, including consultations with businesses citywide on their common gripes, culminating in a report evaluating the "relevance and appropriateness" of the existing regulatory model (N.Y.C. Admin. Code, 2019, p. 9).

The prerogative of the small business owner to be creative and in-your-face with their signage, it seemed, had returned.

Semiotics & technology

In 1972, architects Robert Venturi and Denise Scott Brown launched the hand grenade *Learning from Las Vegas*, which took seriously the design and function of the desert city's over-the-top casino signage and mocked modernist buildings as overblown symbols of cultural erudition. Buildings—especially commercial ones—have denotative meanings courtesy of language on signs, but also connotative meanings that are more suggestive and implicit. They even promoted a binary that can be read as analogous to ours: the "duck" is the pretentious, reference- and art-laden icon, the "decorated shed" is the functional everyman that embraces literalness (p. 87). Their study of the built landscape in two dimensions was revolutionary for a profession obsessed with 3-D space.

For as long as the technology has been available, natural and physical scientists have been utilizing satellite imagery to analyze, for example, deforestation in the Amazon (Sezen, 2021) or destruction of the built environment in wartime ("Satellite images show how receiving aid in Gaza became so deadly," 2025). From the perspective of an average consumer of science and technology news like myself, more traditional, ground-level photography has been less utilized in pursuit of scholarly inquiry.

For urban issues especially, the advent of ubiquitous and free imagery resources like Google Streetview is potentially ripe for exploration. One project that has done this masterfully is media artist Yufeng Zhao's AllText.nyc. Zhao trained an image recognition model to extract visible text from the entirety of Google Streetview's New York City photo corpus, to map, for example, Spanish-speaking enclaves using the word 'iglesia' and paranoid, territorial ones with the phrase 'Beware of dog' (Daniels & Zhao, 2025).

The exact mechanics of how computers "read" images through numeric tallies of color on pixels is beyond the scope of this project. But in *Distant Viewing: Computational Exploration of Digital Images*, Taylor Arnold and Lauren Tilton remind us that "how information is encoded and decoded is shaped by cultural ideologies, embodied ways of knowing, social scripts, and grammars of everyday life" (p. 15). Contrary to what may seem like an objective exercise, the algorithms deployed to train a computer to interpret images were designed by flesh-and-blood humans, with all their social, political and cultural biases.

This was more or less the subject of the class DATA 78000 Topics: Data Bias- How Big Data Increases Inequality. The selected readings for that class collectively argued that data science is a social science, and by imbuing computers with objectivity, we are blinding ourselves to the human prejudices they reflect and inevitably reproduce. Image recognition technology is no different of course, and carries the obvious added baggage of dealing with skin color and facial features.

In one example cited by Ruha Benjamin in her 2020 treatise *Race After Technology*, an AI health and beauty app claiming not only to be unbiased but also promising anti-aging solutions, overwhelmingly favored white faces in its user-submitted beauty contest, having clearly been trained on white standards of beauty. And as Benjamin observes in the same pages, the colorblind utopianism broadcast in the now-famous 1993 *New Yorker* cartoon, "On the internet, nobody knows you're a dog" has proven short-lived by the heavily visual nature of the 21st century internet (p. 43).

DATA SOURCE

I learned about Live XYZ, the source of the storefront imagery, when the founder, Jason Greenwald, gave a demonstration as part of New York City Open Data week. I was so impressed I immediately requested a license to access the full dataset, which was granted as part of a data-sharing agreement the company has with the City of New York (I work at the Department of Housing Preservation & Development). Even though I couldn't immediately think of any use cases that might directly support my work there, I knew it could be a great resource for anyone curious about the everchanging street life of New York City. (I have pledged to Live XYZ that I will incorporate this project into an HPD workflow, perhaps via technology education, to fulfill the data sharing agreement.)

Live XYZ acts as a sort of Google Maps/Streetview on steroids, but specific to the five boroughs. It utilizes a team of human scouts to ground-verify and document the contact information, statuses, and—crucially for me—images of the 100,000+ street-facing goods-suppliers and service-providers in New York City (its ambit is wider than retail, as it includes schools and churches and any other organization or business that has a semi-accessible presence on the street level).

Through their partnership with the City and collaborations with the press, their data have been plumbed to measure the health of physical retail in the lockdown era following the Covid-19 outbreak (Meredith & Sample, 2021) and the phenomenon of landlords in high-value neighborhoods intentionally keeping retail space vacant (Chen, 2024). But as far as I could tell, none had made use of the storefront photography as a subject of analysis.

A dataset like this inevitably has shortcomings in the context of sociological research however, which I began to consider in the context of a close reading of a text called *Data*

Feminism as part of the class DATA 70500: Working with Data: Fundamentals. “Who makes maps and who gets mapped?” ask D’Ignazio and Klein (2020, p. 52). The minority of those of us doing data collection and manipulation almost always have more resources and time than those of our subjects.

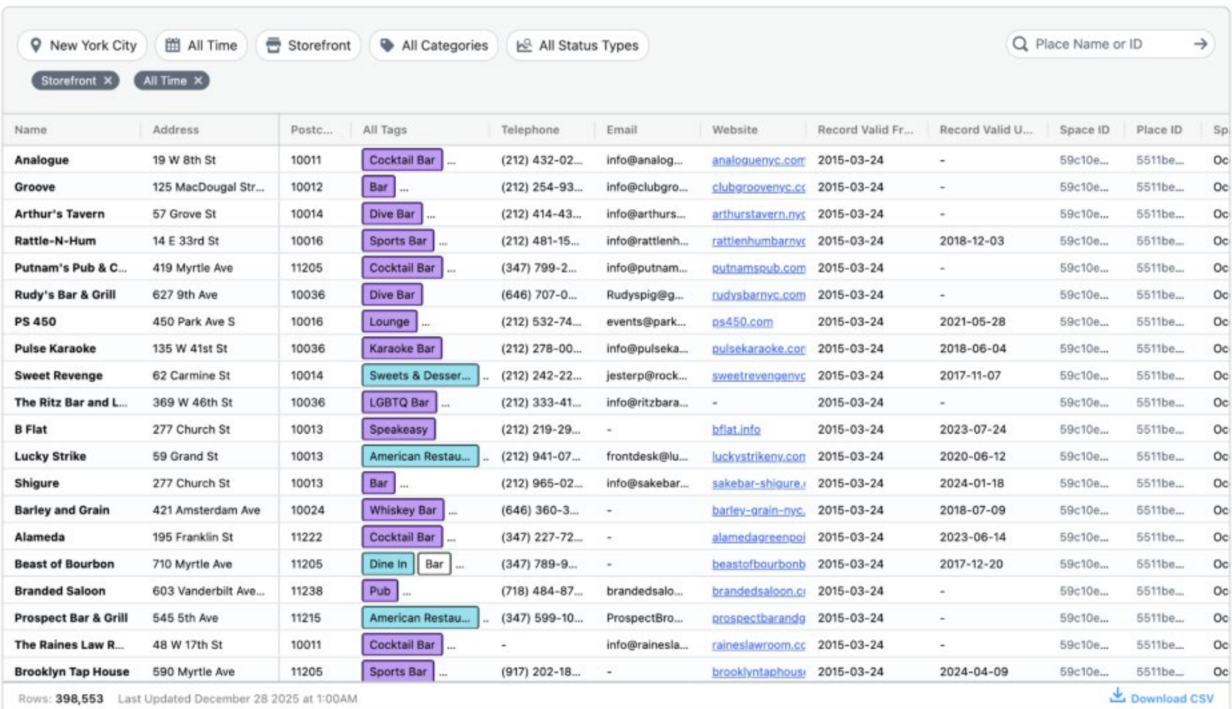
Starting and maintaining a store or a restaurant is a huge expenditure of time, money, and physical and emotional labor. This exercise should aim to not single out for judgment any individual shop-owner or category of business. Rather it should look at the big picture and overall observable trends. For example, should the project encourage readers to cast aspersions on a wine bar because it opened in a surprising neighborhood? I hope not.

Further, Live XYZ is not a public dataset. But by soliciting collaboration and new uses for data and by partnering with municipal agencies, Live XYZ can expect that their data will be used, if not to wholly challenge existing power structures, then at least to further civic goals we expect bureaucrats to be concerned with: eliminating food desserts, creating safe and vibrant communities, or nurturing small businesses to name just a few.

Lastly, this exercise runs the risk of “falling into the wrong hands,” or being abused for profit-seeking. Real estate developers have an enormous interest in investing in the next “hot” neighborhood. So might the findings turbocharge neighborhood change? Of course developers have their own tools to monitor area characteristics and inform investment decisions. But an academic exercise like this runs the risk of *enabling* the very power sources it hopes to challenge.

Business identification

In Live XYZ's backend data preview and download environment, the filters aren't very useful, so I decided to download the entire dataset for all time (existing and closed businesses)—nearly 400,000 records! That download took place on September 13, 2025, so all data used in this project is current as of that date. It comes in a comma-separated value (.CSV) format.



Name	Address	Postc...	All Tags	Telephone	Email	Website	Record Valid Fr...	Record Valid U...	Space ID	Place ID	Sp
Analogue	19 W 8th St	10011	Cocktail Bar ...	(212) 432-02...	info@analog...	analoguenyc.com	2015-03-24	-	59c10e...	5511be...	Oc
Groove	125 MacDougal Str...	10012	Bar ...	(212) 254-93...	info@clubgro...	clubgroovenyc.cc	2015-03-24	-	59c10e...	5511be...	Oc
Arthur's Tavern	57 Grove St	10014	Dive Bar ...	(212) 414-43...	info@arthurs...	arthurstavern.nyc	2015-03-24	-	59c10e...	5511be...	Oc
Rattle-N-Hum	14 E 33rd St	10016	Sports Bar ...	(212) 481-15...	info@rattlenh...	rattlenhumbarnyc	2015-03-24	2018-12-03	59c10e...	5511be...	Oc
Putnam's Pub & C...	419 Myrtle Ave	11205	Cocktail Bar ...	(347) 799-2...	info@putnam...	putnamspub.com	2015-03-24	-	59c10e...	5511be...	Oc
Rudy's Bar & Grill	627 9th Ave	10036	Dive Bar	(646) 707-0...	Rudyspig@g...	rudysbarnyc.com	2015-03-24	-	59c10e...	5511be...	Oc
PS 450	450 Park Ave S	10016	Lounge ...	(212) 532-74...	events@park...	ps450.com	2015-03-24	2021-05-28	59c10e...	5511be...	Oc
Pulse Karaoke	135 W 41st St	10036	Karaoke Bar	(212) 278-00...	info@pulseka...	pulsekaraoke.cor	2015-03-24	2018-06-04	59c10e...	5511be...	Oc
Sweet Revenge	62 Carmine St	10014	Sweets & Desser... ..	(212) 242-22...	jesterp@rock...	sweetrevengenyc	2015-03-24	2017-11-07	59c10e...	5511be...	Oc
The Ritz Bar and L...	369 W 46th St	10036	LGBTQ Bar ...	(212) 333-41...	info@ritzbara...	-	2015-03-24	-	59c10e...	5511be...	Oc
B Flat	277 Church St	10013	Speakeasy	(212) 219-29...	-	bflat.info	2015-03-24	2023-07-24	59c10e...	5511be...	Oc
Lucky Strike	59 Grand St	10013	American Restau... ..	(212) 941-07...	frontdesk@lu...	luckystrikeny.con	2015-03-24	2020-06-12	59c10e...	5511be...	Oc
Shigure	277 Church St	10013	Bar ...	(212) 965-02...	info@sakebar...	sakebar-shigure.	2015-03-24	2024-01-18	59c10e...	5511be...	Oc
Barley and Grain	421 Amsterdam Ave	10024	Whiskey Bar ...	(646) 360-3...	-	barley-grain-nyc.	2015-03-24	2018-07-09	59c10e...	5511be...	Oc
Alameda	195 Franklin St	11222	Cocktail Bar ...	(347) 227-72...	-	alamedagreenpoi	2015-03-24	2023-06-14	59c10e...	5511be...	Oc
Beast of Bourbon	710 Myrtle Ave	11205	Dine In Bar ...	(347) 789-9...	-	beastofbourbonb	2015-03-24	2017-12-20	59c10e...	5511be...	Oc
Branded Saloon	603 Vanderbilt Ave...	11238	Pub ...	(718) 484-87...	brandedsalo...	brandedsaloon.c	2015-03-24	-	59c10e...	5511be...	Oc
Prospect Bar & Grill	545 5th Ave	11215	American Restau... ..	(347) 599-10...	ProspectBro...	prospectbarandg	2015-03-24	-	59c10e...	5511be...	Oc
The Raines Law R...	48 W 17th St	10011	Cocktail Bar ...	-	info@rainesla...	raineslawroom.cc	2015-03-24	-	59c10e...	5511be...	Oc
Brooklyn Tap House	590 Myrtle Ave	11205	Sports Bar ...	(917) 202-18...	-	brooklyntaphous	2015-03-24	2024-04-09	59c10e...	5511be...	Oc

Rows: 398,553 Last Updated December 28 2025 at 1:00AM Download CSV

Figure 3: Live XYZ backend data list preview (Live XYZ)

I then imported that file into a GIS product called ArcGIS. This is the same software we used as part of SOC 81900: Sociological Statistics (Spatial Demography), and which powered my retail turnover project referenced earlier. In ArcGIS, I performed what's called a "clip" to identify 8,639 businesses—still past and present—that were located within the bounds of Brooklyn's Community District 3.[1]

[1] New York City's 59 Community Districts (also called Community Boards) are relatively unchanging, and are a good statistical boundary to proxy for many of the city's ill-defined "neighborhoods." The shapefiles can be found

I then extracted a spreadsheet file containing those location-specific records to transform in Microsoft Excel. I filtered out unoccupied and demolished store locations to leave 6,950 remaining occupied. I also filtered out "temporarily closed," "unsure if operating," "coming soon," "moving soon" and "permanently closing soon" to give just "permanently closed" and currently "operating" businesses. I removed duplicates that were due to a single business going through multiple statuses to end up with the most current status, leaving 4,194 (2,293 operating, 1,901 closed) and moved these to a new tab, 'Closed-operating only.'

I added a column incorporating the logic of: "if Place Creation Date > Space Creation Date, business is newish within the past ~8 years that Live XYZ has been collecting the data."[2] Of 2,293 operating, 1,121 were existing when Live XYZ surveyed (variable times), and 1,172 were introduced *after* Live XYZ first surveyed the location (newly opened business).

I also decided to remove chain stores (indicated in the 'Chain name' column), leaving 3,970 independent shops of the 4,194 total. Even though some are local/small chains, large national chains have fairly strict signage requirements and operate on widespread visual recognition. (This is also an interesting urban and semiotic phenomenon, but is outside the scope of this project.)

Only 224 chain stores out of over 4,000?! For an interesting and scrupulous cataloguing of the rise and apparent fall of chain stores in New York City, see the Center for an Urban Future's annual "State of the Chains Report" (Klein et al, 2025). The chain tagging wasn't perfect however, so I manually filtered out the ones I came across while associating images with

on the website of the Department of City Planning: https://www.nyc.gov/content/planning/pages/resources/datasets/community-districts

[2] Place Creation Date is the date the occupying business/tenant was logged in the Live XYZ database; Space Creation Date is the date the physical retail space/address was logged. See Appendix for the full Live XYZ data dictionary and column descriptions.

each entry (more on that soon). I pulled Boost Mobile, Cricket, MetroPCS, Sprint, T-Mobile and Papa Johns.[3]

Due to time constraints, I decided to focus on new and closed businesses only (leaving out existing but longstanding ones) and hypothesized there would be a greater distinction between these two groups, while the longstanding ones might be somewhere in the middle or "mixed" in the aesthetic binary. So I "hid" the "existing" stores in the spreadsheet, leaving 2,938 of just new or closed. I also deleted entries that lacked a name or were clearly test/dummy data for use by Live XYZ internally.

Photo harvesting

Now that I had my pool of ~2,900 businesses to analyze, I would also need to get each corresponding image. The Live XYZ data pull comes with a column that links to the appropriate business listing on Live XYZ's website ('Live XYZ URL' in the Data Dictionary). From there you can click an image icon to pull up all the photos tagged for the business. In many cases there are more than one, with some interior shots and even close-ups of food.

[3] These mobile phone stores might skirt Live XYZ's "chain" definition by being franchisees, but their signage still fits the visual universality characteristic of national chains.

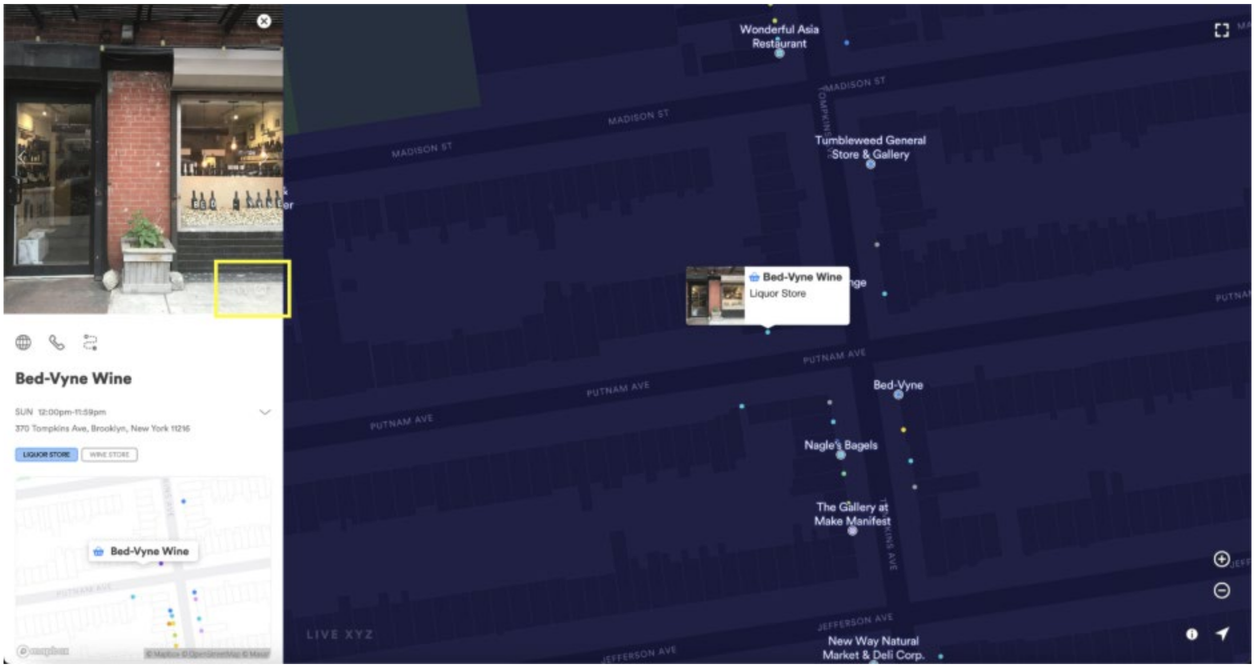


Figure 4: Typical public-facing Live XYZ listing page with image icon highlighted in yellow box (Live XYZ)

Originally, I had planned to send Live XYZ a list of the businesses I wanted photos of, to get them sent in one big batch. But not only was that challenging from an infrastructure standpoint, but the multiplicity of photos for each business meant I wouldn't know if I was getting the "right" one for each. For example, even if a listing had multiple exterior shots, there was often a superior one that showed the whole storefront from a clear angle without much else in the frame to distract (Figure 5). So I ended up manually downloading the best photo I could find for each of the nearly 3,000 businesses!

In addition to downloading hard copies for feeding into the model, I also had to grab the direct photo links from the website and paste them in a new column in my spreadsheet, so that I could include thumbnails of the photos in my website visualizations. Some of the business entries were missing photos entirely, or (based on my subjective judgment) had poor or unrepresentative photos, so I ended up with photos and links for 2,811 businesses (1,701 closed and 1,110 newly opened).

Figure 5: Good example of a close-up, complete storefront image with little else in frame (Live XYZ)

But as I was clicking through so many images, a few other issues came to mind. So many photos were technically correct, but were visited at an odd time so had one of those heavy metal security gates pulled over them; that would certainly affect the image-reading tool. Also, Live XYZ's definition of a storefront is so broad to include anything that's accessible from the street, so places like public schools were included. Since they operate on a much different opening/closing and advertising/publicity logic, is it appropriate to include those? I'll get to some of these questions in the 'Analysis' section below.

MODEL TRAINING/TESTING

Choosing the model

Before choosing the image classification model I would ultimately use, I had to research what was out there and compare them. Some commercial products, including Roboflow and Mistral, looked easy to use and fairly customizable, but would have been expensive to run on the volume of images I had. It was in this round of research that I came across Google's Teachable Machine (https://teachablemachine.withgoogle.com/), which also looked easy to use, and was open-source, so I could download the trained model and use it locally.

Around this time I also started speaking with Stefano Morello, Assistant Director for Digital Projects at the American Social History Project/Center for Media and Learning at the CUNY Graduate Center. He advised that I could go with a pre-packaged set-up like Teachable Machine, or create a custom one, over which I could have more control, but would likely take much longer. LM Studio is an example of a service that lets you run large, complex predictions locally on your computer. Morello offered use of some fast-processing computers at the Graduate Center.

In the interest of time, I leaned toward a faster, simpler approach like Teachable Machine. Its openness made it fast and easy to test with a small initial batch of photos. I ran a test using 10 photos that I had deemed 'old-school,' tagging them as Class 1 in my model, and 10 photos I called 'new-school,' labeling them Class 2. All test and training photos also came from Live XYZ, but without an overlap with the ultimate pool of subject photos. Once trained, the model is able to make a prediction (with between 0 and 100% confidence) about whether any third photos introduced fit which class, and it seemed to be pretty accurate (based on my subjective understanding of the aesthetic binary)! More on the technicalities behind the machine learning to follow.

Another resource Morello turned me onto was a site called Distant Viewing Lab (https://distantviewing.org/). The lab is the outgrowth of a book by the same title from Taylor Arnold and Lauren Tilton, University of Richmond professors of data science and communications, respectively. They seek to break down how computers "see" image components and can make large-scale conclusions about the data found in them (Arnold & Tilton, 2023).

Some of the open-source sample models available on the site looked really promising, but ultimately didn't produce expected results. For example, an "image classification" model is supposed to provide—with a percent confidence—possible objects in a photo (so, for example, a sign or a window, in my test cases), but it wasn't very accurate or likely wasn't trained on sufficient a corpus of images. Another called "image segmentation," which splits images up into object components and identifies the percent of the frame they fill felt like it could be useful for my purposes but didn't seem to be able to finish loading. I'll discuss more in the 'Future directions' section below.

Training 'Teachable Machine'

For the real, non-test model I was going to train on Teachable Machine and then have process my 2,800 storefront images, I collected a training set of 100 photos—50 classed 'old-school' and 50 'new-school.' I sourced these from a combination of businesses referenced in Trinch & Snajdr, businesses I knew personally and could pull up on Live XYZ, and businesses I found just walking around Brooklyn. I made sure these were from outside of my study neighborhood of Bed-Stuy, but representative of a number of diverse neighborhoods at varying stages of gentrification. I have made all the photos used to train the model available here: https://docs.google.com/spreadsheets/d/1PSPU3ZG-4tdYZpHq_FaGankImTOYDALRK2b2FgHT6gU/edit?usp=sharing and in a .CSV file included in the 'Digital Manifest.'

The model was trained using the default parameters set by Teachable Machine: 50 epochs (every sample will be fed through the model at least 50 times) and a batch size of 16 (the number of samples that fit into each round of testing before an epoch is complete). Teachable Machine always splits the samples into 85% training data and 15% test data, to test itself after training itself, all part of the "training" for our purposes. Figures 6 and 8 below show how the model performed on its training and test data, before and after I made an adjustment to the training images based on what I suspected was a confounding factor.

In *Distant Viewing*, the book that informed the image processing model I mentioned above, Arnold and Tilton explain that "there is no need to enumerate why two images are similar" when training a model using buckets of similarity (2023, p. 44). Through the simple ordering of colored pixels, the model can pick up on patterns of both content and composition of the photos it is fed. This is important in my data, as the two storefront styles are somewhat "you-know-it-when-you-see-it," and a human-led annotation of the components that make up each class would be untenably laborious.

After training the model, which takes just a few seconds, I tested a few examples one-off in the browser to make sure it was predicting as well as I had seen in my test run. I then exported the model via TensorFlow (another open-source Google tool for machine-learning), in Keras format (a user-friendly application programming interface), to batch-process in Python. With much help from Cursor (an AI coding tool, more on that in the 'Results/interpretations' section), I then created a Python script to resize the images to 224x224 pixels (Teachable Machine's default) and normalized pixel values to [-1, 1] so my dataset could be run through the model more efficiently (Brownlee, 2019).

However I was still getting results that didn't match my perception of the data. For example, a pretty clear majority of both newly opened and closed business photos were getting pretty confidently classed in the "new-school" camp, despite the pretty obvious minority of this style observed when downloading the images. Per some back-and-forth with Cursor, I was prompted to do some additional conditioning of the data before feeding it into the model: pulling the orientation metadata from the photos so they're all oriented the proper way and resizing the photos with bilinear interpolation, which is a faster way to process this volume of images.

This version of the model started giving more expected results, and that's when I started building the demonstration website around the old- and new-school predictions of the images. The model had produced a value between 0 and 100% confident for each class, and I appended those results to my original CSV file of the storefronts. See 'Results' and 'Analysis & visualization' sections below for more.

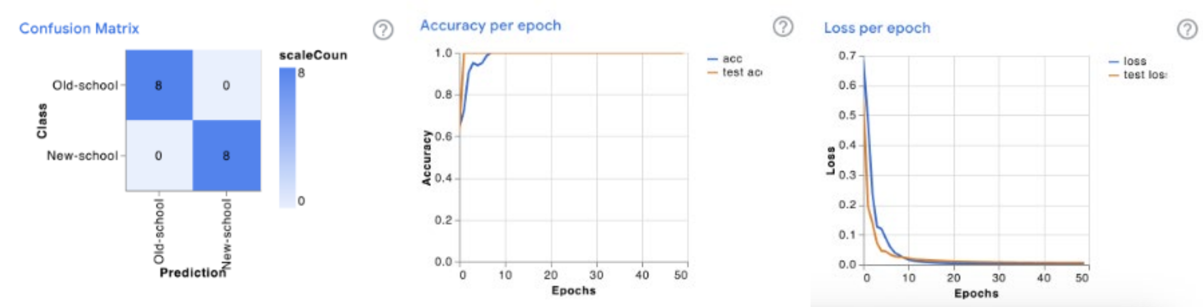


Figure 6: Accuracy and loss (lack of confidence) on the original, binary model (Google)

Looking "under the hood" of the model in Teachable Machine, you can see that both the 85% training images and the 15% test images (the 8 in the confusion matrix) were both increasingly accurate and confident (inverse of 'loss'), the more epochs they ran through. This is the model I would base my analysis and visualizations on.

But I noticed a confounding factor that I thought might be influencing the results. Obviously there would be a wide range of ambiguity between the theoretically perfect "new-school" and "old-school" prototypes. But thanks to messy real-world conditions, both the training data and the sample data employed a mix of typicity, so I was confident the model could handle that. What was more confusing was the preponderance of storefronts that were photographed with those large metal security gates covering them, ostensibly caught at closing hours (Figure 7).

Figure 7: Store photographed with its security gate obscuring key aesthetic details (Live XYZ)

I was worried the model might mistake their brute simplicity for the pared-down new-school aesthetic. So I trained a second model based on the original 100 training images but with a third category of 20 images of storefronts covered by security gates to see if that had an effect on the new/old-school results. The resulting model however, had much less accuracy and confidence (Figure 8). More in the ‘Results’ section following.

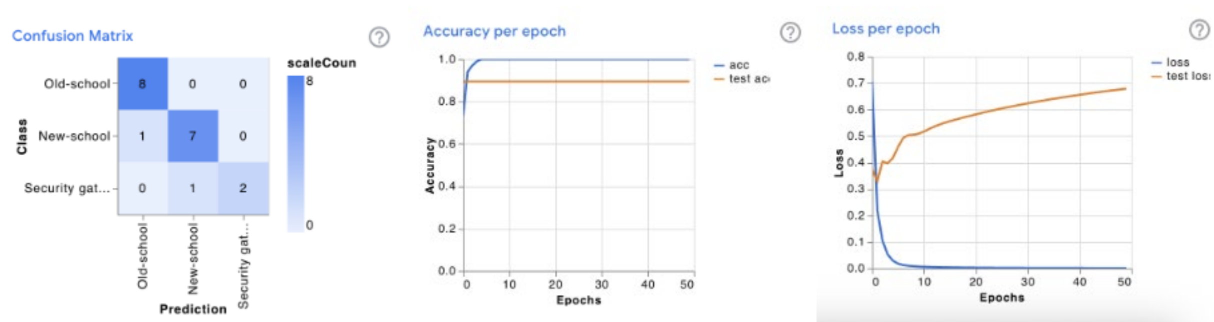


Figure 8: Accuracy and loss on the second, tri-partite model (Google)

RESULTS

In downloading the nearly 3,000 images for the model, I noted some anecdotal visual and thematic trends among the two sets of businesses—newly opened and permanently closed—that I suspected might have an impact on the model's eventual reading. In the new set, since Live XYZ seems to make a point of documenting new storefronts as soon as possible, lots of new places—typically in the "old style" are adorned with temporary flags and vinyl "grand opening" signs.

In the closed set, I observed that lots of the storefronts we wouldn't consider retail (e.g. purpose-built churches, schools, industrial uses like salvage yards) don't fit into the A/B model I had trained the model on. (See the 'Data source' section above for Live XYZ's universe of storefront documentation.) But I figured that the closings of these businesses (there were very few new openings of this sort) still speaks appropriately to the upgrading intensiveness of land use. I also noted lots of closed security gates (which led to my retraining the model), which made me think that perhaps unreliable business hours or slow customer traffic might lead to eventual business closure.

Table 1: Contingency table for initial binary model

	Closed	**New**	**Total**
Old-school	950	548	1,498
New-school	741	561	1,302
Total	1,691	1,109	2,800

Table 2: Contingency table for tri-partite model (with security gates)

	Closed	**New**	**Total**
Old-school	537	329	866
New-school	1,003	729	1,732
Security gates	151	51	202
Total	1,691	1,109	2,800

The retrained model was very good at identifying security gates (it ended up classing ~200 photos this way), but the introduction of a third factor seems to have confused its reading of the primary two categories. Against my initial hunch that security gates might strike the model as "new-school," the majority of the 202 reclassed—124—came from the original model's "old-school." Furthermore, and more worryingly, it moved so many photos from "old-school" to "new-school" that it shifted the overall balance of the dataset from majority old-school to majority new-school, which goes against my subjective reading of the overall dataset.

Based on my spot-checking some of these "shifted" photos, the majority seem like they should have stayed classed "old-school." If I had more time, I might try to remove these 200 entries and re-run the remaining ~2,600 photos through the original binary model. But considering how relatively small that subcategory is, and that security gates tended to be more closely associated with closed businesses, which in turn are more likely to be "old-school," I'm OK with keeping them as part of the big, messy, binary model.

ANALYSIS & VISUALIZATION

Since we have two groups of source material (stores that closed and stores that are new), and two groups of descriptive outcome (old-school and new-school) along a 0-100% probability arc, it's easy to run a regression analysis to better understand the statistical relationship between the two binary categories. These analyses are a product of the statistical methods picked up in DATA 71000: Data Analysis Methods.

The p-value of 0.0002 shows us there is a statistically significant relationship between the new/old-school probability and the new/closed status of the businesses (a 0.02% chance the relationship could be due to chance). But the r-square and multiple R show the predictive power of the model is very slight – only about 0.5% of the closed/open status can be explained by the new/old-school probability of the images. The fairly large sample size of ~2,800 images lets us tease out such a small, yet statistically significant variance.

Table 3: Regression outcome

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.0708							
R Square	0.0050							
Adjusted R Square	0.0047							
Standard Error	0.4880							
Observations	2800							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	3.3564	3.3564	14.0924	0.0002			
Residual	2798	666.4004	0.2382					
Total	2799	669.7568						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*

Intercept	0.3609	0.0131	27.4527	0.0000	0.3351	0.3867	0.3351	0.3867
New-school Probability	0.0756	0.0201	3.7540	0.0002	0.0361	0.1151	0.0361	0.1151

The odds ratio (old-school businesses are 1.31x more likely to be closed than new-school businesses) and the relatively large chi-square statistic (12.3) also both point to the strength of the relationship between status and class. The relatively short timeframe (data collection in this neighborhood began in earnest around 2016) and large volume of store turnover point to a universe of old-school storefronts slowly being replaced by new-school ones.

Visualization

The visualizations provided on this project's website, and described below, were developed with the tenets of color, symbology, and user-friendliness acquired in DATA 73000: Visualization and Design, and some of the interactivity and infrastructure (Observable Plot & Framework: https://observablehq.com/framework/) gleaned from DATA 73200: Interactive Data Visualization. Some of these let us explore trends and relationships that might be hidden in the headline statistics and regression. They are all available here: https://alexmcqw.github.io/Bed-Stuy-signs/

First is the "Old-school vs. New-school" tab. This lets the user scan thumbnail images of the storefront photos and how they were classed by the model: old-school on the left and new-school on the right. They are sorted from top to bottom by the percentage confidence the model had in its prediction. Per the legend, the tiles with dark backgrounds are newly opened, with the lighter ones closed. Users can try to make judgments about what kinds of aesthetics might typify new versus closed businesses, or explore some of the more ambiguous visual cues that might have tipped the model one way or the other.

Next is the "Map" tab, where users can see the businesses of Bed-Stuy (Brooklyn Community District 3) plotted by location. They are color-coded for old- and new-school. As one can imagine, over a period of 8+ years, some locations saw multiple businesses cycle through (up to 5 in a few instances!), so dots for these spaces are enlarged, with the corresponding proportion of storefront style inlaid in a radial diagram. This graphic might be particularly useful for someone familiar with the neighborhood, but I think there are some intriguing insights; it looks like a square near the top left (neighbor to pricier Williamsburg and Clinton Hill) might be a hotspot for the new-school, while the dense corridor of Fulton Street on the bottom left is sticking to an old-school style.

Next, on the "Timeline" tab, users can scroll through history to see each business plotted by opening and closing date. Note: the dates are actually the dates the status was updated by Live XYZ, so there will be a time lag of weeks or months. If you look close enough, you should be able to register slightly more brown-coded (old-school) closed businesses (line moving to the left of the dot) and more pink-coded (new-school) new business openings (line moving the right). This owes a debt to Edward Tufte's notion of "sparklines," the use of visual repetition of a bi-dimensional dataset to make a point in compact space (2010, p. 54).

The last tab has a few more summary statistics duplicative of what was discussed above, followed by a few more exploratory graphics. The first is a stacked area chart showing how the balance of old- and new-school style has gradually shifted in favor of the new over the years with business turnover. Since this combines closed and new businesses at any one time, the total is only ever about half of the ~2,800 photos analyzed. As discussed in the literature review, gentrification is a long, potentially barely discernable phenomenon.

The final graphic shows business turnover at individual address locations for roughly half of storefronts analyzed that shared their space with at least one other over the observed time period. Most of these were just twice over the nearly decade, but six (unlucky?) spaces saw five different businesses cycle through. The weight of the lines connecting the business names indicates a maintenance of the storefront style (faint) or a switch (bold), while the line color reflects the style of the "replacing" business. In most cases, spaces saw a maintenance of the same style between businesses. I suspect this has to do with a landlord's amenability to retail tenant improvements, or, as was vaguely discernible on the map, a propensity for newer style businesses to co-locate (e.g. in the northwestern corner).

I have attempted to structure the website and its visuals as a user-friendly, maybe even fun exploration of aesthetic streetscape measurement and change. As I hope to have demonstrated, the subject of gentrification is deserving of serious scholarly inquiry, and the methodology was appropriately rigorous and quantitative. But in the website, I hope to leave a communications tool that reaches beyond the academy, in the spirit of the "undercommons" as borrowed from the democratizing, subversive concept that undergirded the course DATA 74000: Data, Culture and Society (Harney & Moten, 2013).

LIMITATIONS

The biggest limitation in a project like this is that it's a snapshot in time, in this case September 13, 2025, when the storefront data was downloaded from Live XYZ. But that's also a good reason to continue the project both in time as well in place, either by myself or by others. Even though the old/new-school binary was developed for Brooklyn, and my project focused on an even smaller geography, I suspect these trends are more universal, at least in a United States retail context.

The second obvious limitation was the size of the pool of photos used to train the model. As we saw with the slipping confidence in the second model, a bigger training set would have been the most obvious corrective. However it was fairly time-consuming to identify good prototypes of each class. If I had more time, I would probably work to remove the security gate-obscured data (~200 photos) and then re-run the remaining "good" photos in the original binary model.

But filtering the data in that way brings up another limitation of a project like this—the subjectivity of identifying one class from another, and the wide range of ambiguity that obviously exists between the two extremes. Treating nearly 3,000 small businesses as belonging strictly to one visual camp or another does disservice to all the creativity and chutzpah needed to run an enterprise in an expensive and diverse place like New York City, and broadly paints over divergences in circumstance, personality, capital availability, education, English proficiency, social capital and lots of other human messiness.

As social researchers, it should be our imperative to accompany any finding that relies on "data" with interviews or qualitative, human-generated information; this could be done here by pinpointing a handful of store closures and asking more about the circumstances or trying to track down someone who experienced it. A "visceralization" of this data, per a concept introduced in *Data Feminism*, might focus primarily on the store closures and solicit a sort of grief for what we as a city no longer have, like a *New York* Magazine piece did for restaurants that closed during the pandemic (D'Ignazio and Klein, 2020, p. 84; Swanson, 2020).

In a similar vein, I want to caution against a fatalistic interpretation of the results—that our cities are destined for homogeneity or upscaling, or one that removes all agency or creativity from the Black population of Bed-Stuy. Many of the new businesses that have opened in the neighborhood have been home-grown and these should be celebrated, regardless of how they present. An Instagram account called Black-Owned Brooklyn has chronicled many of these across the borough since 2018 (n.d.).

Finally, the time constraint I gave myself for this project limited the sophistication of the model I could use. As Stefano Morello advised early on, a more customizable model would be more time- and knowledge-intensive to set up, but could be better attuned to specific visual characteristics. That being said, I'm glad the model I chose is open-source and free to use, so others can easily re-create exercises like this.

FUTURE DIRECTIONS

In addition to taking this model and applying it to other geographies, there are a number of directions I think this project could take going forward. If I'd had more time, I would've been interested in bolstering the strength of the end relationship by, perhaps, performing a regression on the predicted class with opening date (regardless of open/closed status for locations where it's available and post-2018), or limiting the universe of observations to food businesses. I think, in addition to their large numbers and the aforementioned cultural sensitivity, they really exemplify this binary better than other categories.

Even though I was unable to make use of the models offered on the Distant Viewing platform, I'm still intrigued by their approach to image processing, namely the 'Image Segmentation' tool. I think this project could benefit from feeding the 'closed' and 'new' photo sets through a model like that and comparing the percentages of components in each group, for example does "text" make up a higher percentage in older storefronts, or does "window" score higher for the newer ones?

A few people suggested making a walk-through-time type of narration, in which I could compare the contemporary storefront imagery with, perhaps Google Streetview, which has documentation in this neighborhood going back to 2011, or even as far back as the publicly available New York City municipal tax record photographs, taken in the 1940s and again in the 1980s (https://nycrecords.access.preservica.com/1940s-tax-photographs/). I think the principal value in seeing something like that is seeing how little the decades-old building stock changes, while superficial storefront signage is much more changeable and subject to trends.

Lastly, I was careful in this analysis to not cast judgment on what kinds of storefront businesses "deserve" to open or stay open in particular neighborhoods. But several of my sources spoke to the grief of witnessing local culture lost to global, capitalist forces, and I think that's an important impulse. So here are some recommendations (theirs, not mine):

- Jeremiah Moss suggests expanding local landmark designation to protect legacy businesses that "contribute to the culture and history of the city" as well as making fines for small businesses lower than for chains (2017, p. 417);
- Moss also referenced a grant program started by the City of San Francisco in 2015 that provides grants as well as red-tape navigation to businesses that have been in operation over 30 years (Toledo, 2025);
- *New York Times* architecture critic Michael Kimmelman argued in 2024 for the idea of landmarking a business not because of any architectural distinction but because of "intangible heritage"; he offered the example of Stonewall Inn whose facade and architectural details are protected but not its actual function as a gay bar. Interfering in the free market may be a tricky proposition, but research around food deserts and the Mamdani campaign promise to develop city-run grocery stores show there's a desire to make neighborhoods healthy, in the most literal sense.

Figure 9: Photo from the dataset of a closed business (that ended up being replaced by an apartment building) with the message "Thank you Bed Stuy" painted on the security gate (Live XYZ)

APPENDIX

Data Dictionary

Live XYZ column headers (utilized fields highlighted)

The column headings in the .CSV file are preceded by "LiveXYZSeptember132025_XYTableToPoint_" which is a remnant of extracting the selected data from ArcGIS. The Field Name (API), second column below, follows that. These first descriptions were provided by Live XYZ. Following, I describe the columns I created.

Common Name	**Field Name (API)**	**Description**	**Format**	**Field Size**
Space ID	spaceId	Unique fixed ID of the underlying physical space which a place occupies. This enables the history of a space to be tracked over time	Object	24
Place ID	placeId	Unique fixed ID associated with each individual business/occupant of a space	Object	24
Name	resolvedName	Name of the business/occupant of a space (or the common/generic name if no name is identifiable)	String	no max
Address	address	Street name and number, along with suite or floor if applicable	String	no max
Postcode	postcode	5-digit postcode	String	5
Space Status	spaceStatus	The status of space during a particular period of time (Values: Occupied, Unoccupied, Construction, Demolished, Unsure if Space)	String	8-15
Place Status	placeStatus	The operating status of a Place during a period of time (Values: Operating, Coming Soon, Moving Soon, Permanently Closed, Permanently Closing Soon, Temp. Closed, Unsure if Operating)	String	9-24
All Tags	tags.name	Names of all tags applied to describe the place. Answers "this place is a ___________"	Array[String]	no max
All Subcategories	subcategories.name	Names of all subcategories that describe the place, abstracted from the tags that exist on the place	Array[String]	no max
All Categories	categories.name	Names of all categories that describe the place, abstrated from the tags that exist on the place	Array[String]	no max
Primary Tag	tagsPrimary.name	The name of the first tag in the list of unique tags that have been attributed to a place. Corresponds	String	no max

		to the tag that most accurately describes that place		
Primary Subcategory	subcategoriesPrimary.name	Name of the category that most accurately describes the place, abstracted from the first tag in the array specific tags that exist on the place	String	no max
Primary Category	categoriesPrimary.name	Name of the category that most accurately describes the place, abstracted from the first tag in the array specific tags that exist on the place	String	no max
Offerings	offerings.positive.name	Names of all offerings, amenities, policies, restrictions, etc. assigned to the place	Array[String]	no max
Hours	hours	The listed operating hours of the business/occupant	Array[String]	no max
Telephone	tel	The public-facing ten digit phone number of the business/occupant	String	10-14
Email	emails	The public-facing email address of the business/occupant	Array[String]	no max
Website	urls.website	URLs of a business/occupant's website(s)	Array[String]	no max
Facebook	urls.facebook	URLs of a business/occupant's official Facebook page(s)	Array[String]	no max
Instagram	urls.instagram	URLs of a business/occupant's official Instagram page(s).	Array[String]	no max
Website Menu URL	urls.menu	URLs of web pages where the business/occupant hosts menus or lists of services	Array[String]	no max
Website Contact URL	urls.contact	URLs of web pages where the business/occupant hosts a contact form	Array[String]	no max
Chain Website URL	chain.urls.website	URL of the website of a businesess/occupants chain affiliation	Array[String]	no max
Storefront	isStorefront	Storefront spaces are permanently fixed at a given location and accessible directly from the street or a public outdoor area	Boolean	4-5
Interior Space	isInterior	Interior public spaces are permanently fixed at a given location, but only accessible from inside another space	Boolean	4-5
Mobile Business	isMobile	Describes whether or not a space is flagged as a mobile space, a non-permanent standalone space or a permanently-fixed standalone kiosk	Boolean	4-5
Parent Name	parent.name	If the place is entered from within or is part of a larger parent space (e.g. mall, park, food hall, etc.), this is the name of the place associated with the parent space	String	no max

Parent Primary Category	parent.subcategoriesPrimary.name	If the place is entered from within or is part of a larger parent space (e.g. mall, park, food hall, etc.), this is name of the subcategory that most accurately describes the parent place	String	no max
Entrance Level	entrances.main.floor	The floor number on which the main entrance to a space is located	Integer	Up to 3
Latitude of Main Entrance	entrances.main.lat	Latitude coordinates of the main entrance to a space, mapped in person	Float64	Up to 18
Longitude of Main Entrance	entrances.main.lon	Longitiude coordinates of the main entrance to a space, mapped in person	Float64	Up to 18
Entrance Method	entrances.main.entranceMethod	How a person enters or engages a space (e.g. through a fixed entryway, walking up to a counter, etc); Values (fixed, walkUp, anyDirection, driveThru)	String	5-12
Short Description	slug	Short-form descriptions, curated from language used on the business/occupant's website or social media pages	String	Up to 200
Long Description	description	Long-form descriptions, curated from language used on the business/occupant's website or social media pages	String	Up to 1500
Live XYZ URL	urls.liveWeb	URL to the profile on the Live XYZ's webmap	String	56
Record Valid From	validityTime.start	The beginning of the range of time over which a state's spaceStatus and placeStatus are valid	Timestamp	20
Record Valid Until	validityTime.end	The end time of a state's spaceStatus and placeStatus. validityTime.end is null when the state is valid at the time of data retrieval	Timestamp	20
Place Creation Date	placeCreationDate	Timestamp of when the business/occupant was created in the Live XYZ database	Timestamp	20
Space Creation Date	spaceCreationDate	Timestamp of when the space was created in the Live XYZ database	Timestamp	20
Last Verified Time	verifiedTimesLast	Timestamp of when the space/place was verified by Live XYZ	Timestamp	20
Official Name	name	Name of the business/occupant of a space as presented to the punblic	String	no max
Generic Name	genericName.name	A generic name given to business/occupant's with no identifiable name based on its classification	String	no max

Chain Name	chain.name	Name of a business/occupant's chain affiliation (if applicable). A chain is a collection of places that share a brand and central management	String	no max
State ID	stateId	The state collects all properties and assigns them a unique ID	Object	24
Generic Name ID	genericName.id	Unique ID relating to field entity	Object	24
All Tags IDs	tags.id	Unique ID relating to field entity	Object	24
All Subcategories IDs	subcategories.id	Unique ID relating to field entity	Object	24
All Categories IDs	categories.id	Unique ID relating to field entity	Object	24
Primary Tag ID	tagsPrimary.id	Unique ID relating to field entity	Object	24
Primary Subcategory ID	subcategoriesPrimary.id	Unique ID relating to field entity	Object	24
Primary Category ID	categoriesPrimary.id	Unique ID relating to field entity	Object	24
Offerings IDs	offerings.positive.id	Unique ID relating to field entity	Object	24
Parent ID	parent.placeId	Unique ID relating to field entity	Object	24
Parent Subcategory ID	parent.subcategoriesPrimary.id	Unique ID relating to field entity	Object	24
Chain Name ID	chain.chainId	Unique ID relating to field entity	Object	24
Primary Photo ID	media.primary.mediaId	Unique IDs of all front portait, exterior and interior images captured in-person by Live XYZ	Object	24

My added column headers

Name	Description	Format	Field Size
Photo_URL	The URL of the selected photo for each business in my filtered subcategories (see 'Business identification' section)	Array[String]	no max
Status_simplified	My tripartite definition of the store's status (closed, new, existing).	String	no max

	See 'Business identification' section for logic. Only 'closed' and 'new' statuses were used for this project.		
Predicted Class	The binary class of storefront style provided by the Teachable Machine model (Old-school vs. New-school)	String	no max
Prediction Confidence	The confidence the model had in its 'Predicted class' prediction (between 0 and 100%)	Decimal	(0-1) up to 4 decimal places
Prediction Source	The source folder the associated business's image was found in ('Closed' or 'new'); should match the 'Status_simplified' column	String	no max

REFERENCES

Arnold, T., & Tilton, L. (2023). *Distant viewing: Computational exploration of Digital Images*. The MIT Press.

Barnett, D. (2025). *Welcome to 2465 Atlantic Ave Brooklyn, NY; Noble; N.Y.S.M.* Brooklyn, NY; New York Sign Museum.

Benjamin, R. (2020). *Race after technology: Abolitionist Tools for the new jim code*. Polity.

Black-Owned Brooklyn [@blackownedbklyn]. (n.d.). *Posts* [Instagram Profile]. Instagram. Retrieved January 6, 2026, from https://www.instagram.com/blackownedbklyn/

Bradley-Smith, A. (2024, June 7). Neighbors organize against deed theft, call out investors who target heirs. *Brownstoner*. Retrieved January 5, 2026, from https://www.brownstoner.com/real-estate-market/deed-theft-rally-bed-stuy-212-jefferson-avenue/.

Brownlee, J. (2019, July 5). How to Manually Scale Image Pixel Data for Deep Learning. *Machine Learning Mastery*. December 28, 2025, https://machinelearningmastery.com/how-to-manually-scale-image-pixel-data-for-deep-learning/

Chen, S. (2024). The zombie pharmacies that are holding back new york city retail. *New York Times*. Retrieved December 31, 2025, from https://www.nytimes.com/2024/08/06/nyregion/pharmacies-vacant-drugstores-retail.html?searchResultPosition=1.

Cheung, A. (2017, October 30). Anti-gentrifiers say new pilsen restaurant puts “our lives ... in danger.” *DNAinfo*. Retrieved December 30, 2025, from https://www.dnainfo.com/chicago/20171027/pilsen/anti-gentrifiers-say-new-pilsen-restaurant-puts-our-lives-danger/.

Chideya, F. (2006, July 20). 'There Goes the "Hood": Gentrification in New York City. broadcast, NPR.

Coleman, J. (2016, May 3). How these wooden fences became a symbol of gentrification across los angeles. *LAist*. Retrieved January 3, 2026, from https://laist.com/news/wooden-slat-fences.

Cortright, J. (2025, March 5). Our 2025 list from A to Z of everything that causes gentrification. *City Observatory*. 2025, https://cityobservatory.org/everything-that-causes-gentrification-from-a-to-zv2-2/

Daniels, M., & Zhao, Y. (2025, July). *NYC's urban textscape*. The Pudding. https://pudding.cool/2025/07/street-view/

Dawid, I. (2014, January 26). Gentrification and Displacement: Not the Relationship You Might Have Thought. *Planetizen*. December 29, 2025, https://www.planetizen.com/node/67079

D'Ignazio, Catherine, and Lauren F. Klein. *Data feminism*. Cambridge, MA: The MIT Press, 2020.

Freeman, L. (2006). *There goes the 'hood: Views of gentrification from ground up*. Temple University Press.

Garza, F. (2023, May 16). The gentrification font: how a sleek typeface became a neighborhood omen. *The Guardian*. Retrieved January 3, 2026, from https://www.theguardian.com/us-news/2023/may/16/neutraface-font-gentrification.

Harney, S., & Moten, F. (2013). *The Undercommons: Fugitive Planning & Black Study*. Autonomedia.

Hipsters of NY [@hipsters_of_ny]. (2025, May 16). *"R.I.P. Astoria"* [Photograph]. Instagram. https://www.instagram.com/p/DJuEpiHAgMd/

Kimmelman, M. (2024). How can we save the best parts of our cities? *New York Times*. Retrieved December 30, 2025, from https://www.nytimes.com/2024/11/30/arts/design/preservation-cities-new-york-intangible-heritage.html?smid=nytcore-ios-share.

Klein, S. A., Neches, R., & Bowles, J. (2025). (rep.). *State of the Chains, 2025*. Retrieved December 2025, from https://nycfuture.org/research/state-of-the-chains-2025.

Kode, A. (2025, July 25). *Saving New York's midcentury signs - The New York Times*. The New York Times. https://www.nytimes.com/2025/07/25/realestate/neon-sign-museum-new-york.html

Lang, M. J., & Harden, J. D. (2025, March 2). The house color that tells you when a neighborhood is gentrifying. *The Washington Post*. Retrieved January 3, 2026, from https://www.washingtonpost.com/dc-md-va/interactive/2025/gentrification-gray-changing-city-neighborhoods/.

McQuilkin, A. (2025, November). *Churches out, tattoo parlors in: measuring retail turnover in New York City's "gentrification core."* Manifold @CUNY. https://cuny.manifoldapp.org/read/nyc-retail-gentrification/section/054637fa-f4c2-44be-9ab4-354a8ca240be

Meredith, M., & Sample, H. (2021). *Vacant Spaces NY*. Actar Publishers.

Minority and Women-owned Business Enterprise (M/WBE) Certification Program. Retrieved 2025, from https://nyc-business.nyc.gov/nycbusiness/description/minority-and-womenowned-business-enterprise-certification-program-mwbe.

Moore, L. V., MSPH, & Diez Roux, A. V., MD, PhD (2006). Associations of Neighborhood Characteristics with the Location and Type of Food Stores. American Journal of Public Health, 96(2), 325-330.

Moss, Jeremiah. (2017). *Vanishing New York: How a great city lost its soul.* Dey St.

Murray, J. T., Murray, K. L., & Stein, C. (2023). *Store front NYC: Photographs of the city's independent shops, past and present*. Prestel.

N.Y.C. Admin. Code § 28-213.1 Department penalty for work without a permit. As amended by Local Law 28 of 2019: https://www.nyc.gov/assets/buildings/local_laws/ll28of2019.pdf

New York City Small Business Services. 2012. Storefront Improvements: A Guide for Commercial Districts. New York City: City of New York. https://www.nyc.gov/html/sbs/downloads/pdf/neighborhood_development/storefront-guide.pdf (downloaded via Wayback Machine: https://web.archive.org/web/20210829205448/http://www.nyc.gov/html/sbs/downloads/pdf/neighborhood_development/storefront-guide.pdf)

Reid, N. (2020). Do craft breweries gentrify neighborhoods? it's complicated. *Salon*. Retrieved December 30, 2025, from https://www.salon.com/2020/01/13/are-craft-breweries-a-harbinger-of-gentrification-its-complicated/.

Satellite images show how receiving aid in Gaza became so deadly. (2025, July 28). *The Economist*. Retrieved December 30, 2025, from https://www.economist.com/interactive/graphic-detail/2025/07/28/satellite-images-show-how-receiving-aid-in-gaza-became-so-deadly.

Sezen, U. (2021, December 17). A Satellite-eye View of Three Decades of Deforestation in the Amazon. *Nature Documentaries*. December 30, 2025, https://naturedocumentaries.org/20391/satellite-eye-view-decades-deforestation-amazon/

Swanson, C. (2020). Reasons we've loved new york: a send-off to the many places, big and small, that closed in 2020. *New York Magazine*. Retrieved December 31, 2025, from https://www.curbed.com/article/nyc-businesses-closed-2020-pandemic.html.

Toledo, A. (2025). San Francisco hands out money to legacy businesses — but it couldn't save these shops. San Francisco Chronicle. Retrieved January 10, 2026, from https://www.sfchronicle.com/sf/article/legacy-small-business-registry-20290581.php.

Trinch, S., & Snajdr, E. (2020). *What the Signs Say: Language, Gentrification, and Place-Making in Brooklyn*. Vanderbilt University Press. https://doi.org/10.2307/j.ctv160btqs

Tufte, E. R. (2010). *Beautiful evidence*. Graphics Press.

Urban Dictionary. (n.d.). Gentrifier. In *Urban Dictionary*. Retrieved 2025, from https://www.urbandictionary.com/define.php?term=gentrify.

Venturi, R., Scott Brown, D., & Izenour, S. (1972). *Learning from Las Vegas Robert Venturi, Denise Scott Brown, Steven Izenour*. TPB.